JIU/REP/88/6 (Part I)

对联合国技术合作项目评价系统的评价

第一部分

联合检查组

阿兰·古尔唐编写

日内瓦
1988年

对联合国技术合作项目评价系统的评价

第一部分

联合检查组

阿兰·古尔唐编写

目　　录

	段　次	页　次
导　言	1 - 12	1
一、项目评价与逻辑	13 - 42	3
评价者与被评价者之间的关系	15 - 19	4
项目可评价性的程度和水平	20 - 28	5
从项目角度所看到的限度和水平	29 - 35	7
评价的客观性和功能	36 - 42	8
二、项目方法和国别方法	43 - 77	10
持续性和影响的评定	43 - 55	10
影响协调和竞争	56 - 62	13
国别方法的必要性	63 - 77	14
三、评价程序的改善	78 - 113	17
手册和理论之间的时间滞差	79 - 87	17
程序的及时性	88 - 97	19
各种评价的调整	98 - 107	21
准则的修改	108 - 113	23

导　言

1．本研究报告的目的并不在于审查据称过去几年联合国在某一特定区域所从事或为联合国从事的项目的评价时所出现的一切问题。它并不试图辩明已进行评价的结果如何，也不试图审议可从中得出的结论。其目的范围较有限，但规划较宏伟，即评价联合国范围内所采用的评价系统。因此它处理的是项目而不是方案，而后者是联合检查组另一份报告的主题。它全面审查：

(a) 评价国别项目而不是区域项目的概念、方法和做法；

(b) 评价这些项目所用“准则”的内含哲学和内容以及评价者、被评价者和关于可评价性的推断之间的相互关系；

(c) 监督和评价的结合与／或分离，其各自重要性和包括范围；

(d) 由于愈益依赖捐助国政府执行项目而影响所有评价系统的新束缚；

(e) 为了改进其在联合国范围内的行政受委托进行评价的部门的活动和安排，以及为了改进这些部门在联合国范围内的行政与职能而可能需要的任何改革或改变。

2．为了保证这种研究不仅是纯理论的，而且它将针对不太严格的程序和组织改革提出建议，似乎有必要采取切实可行的办法并从以下几个执行机构的经验中汲取实例加以具体说明：粮农组织、工发组织、教科文组织、儿童基金会和难民署，这些机构与开发计划署合作开展工作；选择它们不仅仅是因为它们发挥了作用，而且事实上还因为它们的职能、目的和办法分殊差异。选择亚洲和太平洋区域，是因为它最大程度地提供了各种各样的经济制度、领土面积、地理状况、社会文化环境和发展水平。

3．报告未载有固定的职权范围。不能强求承担某一任务的联检组的检查专员所开展的调查符合国际公务员依照其等级起草报告所应遵守的程序。检查专员所进行研究的价值不在于遵守某一职权范围，在个人而不是集体行使职权时，他本人完全有资格赋予自己这种职权。这种研究的主要价值在于调查过程中有所发现，检查专员有能力改变其调查的范围或方向，当然不能忘记其既定目标。

4．尽管没有人要求对所有评价系统拥有包罗万象的知识—— 这实际上甚至经验极为丰富的专家也难以做到，但报告所提建议的基础是分析大量行政和学术材料，特别是在现场进行观察和视察项目，并在１９８７年为期一月的出差期间在东南亚和太平洋区域的十几个国家中会见了一百多个接待国政府和专门机构的官员。

5．对本报告初稿的一种批评意见是，它既未大量援引决议或转录决议案文，也未通过某个文件来证明每一观察结果。可以这样说，列举文件会使这种报告篇幅太长，这是毫无疑问的；这对支持一种论点也许是必要的，但如果为了提出切实可行的建议，则无此必要了。研究在较大程度上依靠需适当考虑其保密性的资料来源或谈话，而不是依靠大量的书面调查表，因为这种调查表在多数情况下所得到的答复太多外交辞令或过份基础知识因而无法采用。

6．有些人还认为初稿带有主观性或倾向性。这种批评者忘了这种所谓的“主观性”来自这样的事实：这种讨论是非正式的，而且需要整理。口头讲话比书面文字中往往更容易发现真正客观性，关于客观性的结论很少能呈交总部，而在现场的联合标查组检查专员则比较容易发现这一点。

7．本报告问世之时正值联合国评价系统似乎诞生不久却已充分确定，足以审查这些系统的方法和方向，而不必担心各方怀疑其存在的理由。报告遵循联合检查组采用的原则，特别是检查专员索姆在其１９７９年和１９８５年关于评价问题的报告所采用的原则。因此所提出的建议具有相当普遍的性质，超越了执行机构的分界线。评价不是作为一种孤立的管理工具，而是作为构成被研究机构的一套决策程序的组成部分的一种技巧。

8．过去，人们追求定义过份含糊的目标和努力为业务活动筹措资金并提供投入，对这些活动的重视超过项目的适用性和有效性和其受益人真正利益和性质的考虑等问题现在这种时代似乎已永不复返。促进对评价之有效性、甚至必要性的认识的事件已众所周知，只需要简单地加以回顾。捐助国在“石油冲击”之后所遇到的预算困难使它们更密切地关注其捐款是否用于多边援助。由于提供的资金相对减少，捐助国开始更多地注意减少援助给它们带来的好处。至少其政府经常被敦促考虑这方面问题，而同时一些发展中国家本身已成为净捐助国。

9．现在人们接受这样的事实，即评价是一种正常和有效的项目管理工具，在这方面，开发计划署行政方面发挥了重要作用。联合检查组的许多报告极力确保人们接受评价的概念，普及了其词汇并在执行机构中扩大采用其方法和支助系统。联合国财政危机和旨在增强组织体制的１８人小组的建议使得更有必要求助于原先作为治标剂而此后又成为项目活动必要成分的手段，这是愈益感觉到的需要，是一种关键的行政需要，在某些方面是一种潮流，有时甚至是一种托辞。

10．自１９８１年以来，强调和重申支持执行机构实施和组织评价能力的大会决议不乏其例。１９８１年36／228号决议鼓励各机构为此目的在这方面与联合检查组合作。１９８２年第37／234号决议的附件确定评价是一种允许继续评论和审查成就和与此有关的集体意见的系统。１９８３年第38／227号决议请秘书长审查增强如此确定的系统的一切现有可能办法。大会经由其１９８５年第40／259号决议和１９８６年第41／213号决议更有力地重申它曾反复表达的决心，并特别强调有必要利用联合国的评价能力为各成员国的决策提供比较有事实根据的基础。

11．正是应该从这一角度看待本报告。本报告根本无意同联合检查组以前研究所得结论唱反调，而是开始继续已经展开的关于评价问题的调查。评价并不是一项不可改变的目标，也不是一种一成不变的教条。因此自然应该大大改变其方向，审查其方法，划分其类型，制定新方针。

12．经验表明，评价的主要目的和某些类型的评价的明显必要性随着试验和开始阶段的推移而变得模糊起来。本报告所受到的批评和赞成的程度之强烈不仅证明——如果有此必要—— 所提出问题的重要性和关联性，而且还证明它们所引起的日益明显的动摇不定和焦虑不安必须加以克服。

一．项目评价和逻辑

13．对联合国内部的发展业务的评价是最近历史上的一种变革，与此同时，在社会科学和应用研究以及在政府一级，对待这一进程的处理又出现了显著的变化。首先是一些工业化国家普及这种做法。这必然促进使用新拟定的方法，并影响到联合国系统专门机构执行项目的制定和实施。

14．评价最初采取调查形式。调查依靠比较详细的调查表使得有可能了解受抽样调查者或社会阶层的观点和愿望，然后分析每一回答人对恰当性作出不同程度个人判断的答复。调查无可辩驳地改善了先前资料收集活动，但其有效性却受到怀疑。其突出的缺点是成本高昂，而产生的效果却与支出似乎不成比例。

评价者和被评价者之间的关系

15．这种调查概念的方法似乎太主观，性质上偏重叙述，因此有人认为应该发展到这样一种评价系统的概念：这种系统试图根据组织活动的目标“尽可能系统地和客观地”（JIU/REP/79/2）确定这些活动的关联性、有效性和影响。新方法的目的是合乎科学、侧重数量而不是质量，它强调原因与效果、手段和目的、投入和产出之间的关系。其目标是通过使资源、货物和服务与活动之间、活动与产出之间以及产出与目标之间的顺利相互作用，凭经验监督项目的适当执行。

16．这样，项目被比作一种进程，但其结果却称为产出，人们说它所依靠的逻辑一部分是机械和线性的，一部分是分析性的，而且还具有一定程度的抽象性。有一种假设认为，认真规划和适当管理必然会把投入转成产出。项目被看成是一系列手段和目的，一连串原因和结果、各种等级的层次，这些因素相继出现并相互补充。它表面上是一系列的中间目的，而这些目的转变成手段．又是连续阶段的组成和分解以及成为原因的结果。其中机械性因素优于结构性和人的因素。

17．项目系开发计划署和执行机构通过的准则的主题其技术性、策略性和分门别类性质，相当于泰勒类型的分工。它们作为独立的实体而不是作为一个综合体的组成部门发挥作用。项目似乎象在惰性介质中，在实验室实验中，而不是在实际生活的气氛中发展：项目在形成过程中似乎不是注定要成为社会现实，而社会现实倒似乎是物体；尤其是发展中国家的经济现象似乎没有基本的文化属性。

18．然而，评价可适用于各种主题，而较少用于国家，但它十分经常地可以在一个项目、一组项目或有时在一个方案方面发挥作用，尽管人们总是无法把方案同以平庸方式连结的一套项目区别开来。因此显然，项目的拟订和实施必然会影响评价的进行和成绩。甚至在评价者理解和设想可评价事物之前，他就反映了被评价者。

19．开发计划署手册、有关执行机构手册和亚太经社会手册在确定利用项目管理工具的方法—— 拟定和评估、监督和评价时，着重阐述评价者和可评价事物，而只是附带地提到地点和工作范围可能会产生的差别。按照这些指导原则内含的哲学，评价的有效性与受评价项目的特征既无联系也没依存关系。按照联合国协商一致观点所主张的成员国平等原则和不容置疑的发展必要性，指导原则认为应该从有关多边援助活动的规则的普遍性中产生一种逻辑，以及一种适用评价标准程序的相互关联做法。

项目可评价性的程度和水平

20．当然，指导原则并没忘记考虑到外部因素，这些外部因素的定义是不受项目管理者控制或重大影响的事情、行动或形势，但其成功与否具有重要意义。因此开发计划署项目实绩评价报告的格式中提到这些因素。这些外生变数虽然分成各种类型（体制、政治、社会文化、经济和“其他”），但未引起很多注意。拟订或评估阶段都未提到这些因素，只是在评价的某一阶段才提到，这时项目负责人被要求报告项目产出或叙述与投入有关的内部业务因素。似乎外部因素必须在末期而不是在初期加以考虑，在项目过程中而不是在其开始之际，在危险和结果一级而不是在范围和投入一级。

21．因此“外部因素”只涉及项目实绩评价报告所包括监督和评价的重要问题。虽然附件的表格和工作文件中提到这些因素，但开发计划署关于监督和评价的章节中没有提到。它们只限于项目实绩评价报告范围内，而没有出现在关于深入评价权限的文本中。开发计划署不仅只给它们以有限的地位，而且在无须详尽无遗地列举这些因素的过程中也未提到有些极为重要的因素。这些外部因素的例子有平行的双边援助网、非政府组织同时发生或相辅相成的活动以及执行项目国家的统计和文献机构的质量。

22．从项目拟订和初步审议阶段通过监督直至评价，开发计划署和执行机构手册过度倾向于使项目脱离其社会文化环境、已经取得的发展水平和经济体制的方向。不论基本数据收集是否受到资料储存系统健全与否、行政结构稳固与否或内部联络

有效与否的影响，评价采用同样的效益标准。手册对最不发达国家、发展中国家和新兴工业化国家不加区别，而对这些国家，评价应在不改变其基本意义的条件下至少适用相对不同的评估标准和采用不同的进度指数。

23．然而从１９８８年起将由工发组织适用的“项目拟订指导原则”的编制规定在一定范围内应参照关于项目所属经济部门、东道国发展战略、方案与体制环境和行政基础设施的说明。这一编制方式的优点是从头抓起，而缺点是没有考虑到社会文化环境，或必然会对项目产生影响的人口和政治因素。

24．开发计划署的指导原则也似乎不再述及按照应评价项目的性质相对预先确定任何评价，所以外地解决的可能性没有引起应有的注意。当然，它们按照项目主要目标是否是建立或加强体制结构（研究、规划、教育），提供直接支助（研究、调查、咨询服务）或训练（训练班、讨论会、讲习班、奖学金、研究旅行），不在这些项目之间作功能的、客观的区分。它们还把项目划分为试验或试验项目、援助项目或与投资或投资前有关的项目。它们还规定一种特殊的项目，这些项目在支持国家项目的名义下使捐助国获得行政和后勤支持的手段。对于联合国范围内的职能性方案预算，手册的分类提供了宝贵资料并使得能够进行可以实施的细目。

25．但它没有提供任何足以说明问题而限制性不太大或过分详细的纲领性文件。它没有附带确定相对可测程度的公式或考虑到项目目标，作用和要素的项目可评价性水平的模式。更糟糕的是，这种分类忽视了农业与工业之间、行政与研究之间、文化与健康之间以及职业与社会之间的差别。因此它不可避免地忽略了这些方面和特性的多样性对评价的可行性和效力的影响。

26．手册应更多地注意项目的特殊性质与项目评价之间的关系。人们没有可能始终仿效同一模式，也无法把拥有不同目标、执行过程进度不一或不适应同一方法的各种项目放在同一模型里。因此拟订、监督、初步研究和归纳评价的程度应该因所涉活动和部门而异。

27．简单的例子可以说明这些问题。无论是机场、桥梁、还是通讯，项目需要少量伙伴，彻底的可行性研究已进行，参与工程者的能力已令人了解，除非出现重大失误或特殊的建造错误，项目执行将采取可预料的路线。监督比评价更为重要，

因为评价太晚只能发现失败或任何可能的损害。它的重要性不如拟订和初步研究阶段，因为主要困难出现在筹备阶段，而不是在项目执行阶段。

28．就农业或保健项目而言，许许多多的可能伙伴和受益人，初期难以预见的各种困难使得有必要较少集中于项目拟订，而应更为重视通过实践和监督学会管理综合农村发展项目被恰如其分地比作在无航海图的海洋上航行，航线和航向必须按照发现的情况加以修正。项目拟订和要求确定绝不能予以忽视，但在这种情况下，其重要性不如执行阶段的试验和纠正、保证选择和寻求灵活解决办法。在这种情况下，必须强调监督而不是评价。

从项目所看到的限制和危险

29．项目的性质影响到评价的类型，必须重视拟订和初步研究，详细或粗略的监督和评价、事先评价和事后评价、执行过程中的纠正以及筹备、影响和后续行动。恪守过度标准化的规划而不对项目的特征予以足够注意，这会妨碍任何结论性的项目评价。

30．项目逻辑和评价逻辑以作为其必然结果的个别项目为基础，不重视前后关系而且忽视具体情况，这意味着辅助因素或微不足道的质量可能被看作是一种讨厌的东西或限制。这对确定业务活动成功与否可能是至关紧要的。这一逻辑把评价的范围限于特定项目或一组项目，从而限制了评价者的目标，仅仅使他得到有关需要他评估的项目的歪曲观点。

31．项目执行不论在其评价者眼里如何合适和有效，可能实际上掩盖问题，从而使这些问题的解决更为困难。它使捐助国的注意力集中在特许执行上并优先拨给它可用的国家资源，因而加剧了没有显露但在阴影中徘徊的困难。逐个项目审查的办法会使受益国有关各部之间产生障碍，并使各种措施的资金筹措复杂，而这些措施的成本和影响力还不是规划的主题。一旦项目完成和援助停止，它往往取消其经常开支的资助。但由于评价者必须评价项目的执行，因此评价者不会考虑到优惠较少部门可能已经放弃或损失的东西。

32．当然有人会主张，捐助国政府的计划有连贯性和持续性，而这种计划在原

则上应避免这种扭曲和阻碍。然而，一些发展中国家的计划管理和各部门之间（财政、计划、技术部）的竞争的情况是众所周知的。人们普遍认识到，项目必然会对国家的计划产生扰乱的影响，于是各种捐助国之间发生竞争。这自然会妨碍方案的顺利筹备和项目拟订的和谐发展。

33．从国际组织及其专门工作人员的部门观点来看，项目的执行和资金支付率似乎十分出色。评价者根据表格和调查表会得出结论说项目获得成功。但事实上实际效果可能很有争议，因为有时低估了工作对象的财政、道德和行政资源的筹集。可能没有采取任何措施把受益人同这一工作联系起来，因此其参与发展进程的意愿可能会遭受挫折，因此在一个亚洲国家里，人们注意到社区发展和主动精神的地方传统在农民环境里大为退化，而多边援助在一定程度上使这一环境瘫痪，而不是使之更有生气。

34．如果捐助国和项目管理队伍试图发挥过分突出的作用，并只想减轻受援国的负担，因而取消或看轻其官员和其工作对象的潜在贡献，那就会剥夺他们获得知识与经验、技巧和心理投入的源泉，这些因素历来是无法评估的，至少从项目角度是无法评价的。

35．同样，不仅国家政府而且连地方企业也应参与规划和技术磋商，即使其以成本效益衡量的有效性证明较低。发展不在于通过不同程度地大量输入多边或其他形式的援助来满足受益对象的需要。这是一种社会进程，而不是一种产出。因此这是一种代价，而被项目逻辑禁锢的评价是无法予以评估的。社会事实也是心理事实，而发展一旦偏离项目的逻辑，评价则毫无用武之地。

评价的客观性和功能

36．由于评价的目的是象所评价的项目一样合理，它寻求通过无限扩大其适用范围来提高其客观性的方法。例如，项目的初步研究可作为事先评价，人们听说在案件中既做法官又做当事方的某人作“自我评价”的例子。评价的类型是按照工作的强度、评价者的来历和类型、评价的主题及其所得到的资金数额来划分的。评价也可以是摘要的或回顾的，定期的或专案的。有时“评价”这一词可用在只是一种

“评估”的工作，或只是起这种作用，而采用“方案审计”的题目。

37．由于联合国及其专门机构对于评价的兴趣似乎正在浓厚起来，现在的趋势是，把这一术语用在越来越多与真正评价无甚关系的报告和研究上面。评价在一定程度上成了其自身成功的受害者．似乎既是万灵药，又是行政上的权宜之计。一旦所有事物都自称为评价，这一术语的真正含义就令人怀疑了。这一概念在神话和词典定义之间摇摆，因此它在范围方面得到的好处却失之于精确方面。

38．在这种情况下，要维持评价的假定客观性，就必须接受评价的结论。必须确保有用的资料按时或按预定的时间间隔送交适当人员，确保有关人员收到紧急或重要事项不致太迟，并避免仅仅间接有关的人员陷于多余或次要的资料中；这些要求是简单的常识。但这并不意味着它们因此成为普遍习惯，尽管它们是客观评价的必要补充部分。

39．经审查发现，开发计划署和执行机构的手册似乎没有足够注意保证调查表和表格切实可用。在调查表的框框里填写的内容自动回答了业务人员可能忽视或忘记的问题。这有助于把数据输入总部的计算机。但人们并不总是能够以是或否来回答，或通过在某一个框框里打勾作出合适答复。过分藐视叙述和过分依赖标准表格有可能只记载细微末节。如果没有现场检查（对此很少有什么忧虑）和普遍检查制度（这是迄今机构级所缺乏的），在许多情况下，所收集的答复的有效性可能会令人怀疑。

40．表格的简单标准化对机场入境检查和材料控制来说也许是合理的，但必然会损害客观评价。过分复杂化也应予以避免。１９８６年底，儿童基金会发表了一本用于外地辅助工作人员训练班的“审计指导”。这本长达１５０页的指导书对所有偶发事件作了规定，它似乎理所当然地认为，受到审计的国家拥有有时甚至在高度工业化国家也难以找到的行政和统计设施。

41．这种详尽无遗的指导书采用调查表的形式，必定花很长时间才编写成功。但其作者忘了一件事，即外地工作人员必须耗费几个月的专职工作才能得到正确答案，从而妨碍了其它同样重要的工作。此外，其作者没有足够地注意，使评价手册切实可用。需要确定的主导趋势和主要事实消失在所提出的多种问题中，因为这些

问题不是以优先次序安排的。这一事例尽管是极端的，但不是孤立的，它说明必须急速制止的一种普遍趋势。

42．不尊重主次顺序就不会有客观评价，这意味着必须对各种结论进行有条理的分类。尽管显然有必要得到明确的答复，调查表和表格也必须要求评价者突出显著要点和重要数据。如果只是为了确定重要的因素而使表格过分简单，仅仅记载可以衡量和平常的事物，就可能忽略最值得注意的方面。如果表格过分复杂，会出现相反的可能性。以基本的方式表达一切事物等于什么也没表明。要求说明一切事物是同意一无所述的另一种方式。

二．项目方法和国别方法

持续性和影响的评定

43．开发计划署急于确定它所资助的项目和方案的产出是否实际上在持续，这是情有可原的。开发计划署依据的假定是，项目的“持续性取决于工作对象或机构”连续和有效地“受益于项目结果的能力”，从而得出结论说，这是一个受许多变量影响的一个目标，而这些变量必须在项目拟订和项目实施阶段以及在项目完成之后加以考虑。必须确定项目拟订可能妨碍或有利于这种“持续性”的程度，以及回顾式评估可能如何有益于项目实施的专题评价。另外还必须更好地理解和更彻底地研究项目结果如何和为何继续或停止促进全面发展努力。

44．各执行机构理事会，特别是粮农组织理事会越来越多地考查项目“持续性”或可行性问题。从这种“持续性”考查中得到的第一个教训似乎已证明相当有用。尽管实际结果还有待于确定。这些教训表明，应在项目筹备和拟订阶段而不是在实施阶段强调“持续性”问题。例如，就农业项目而言，生产、销售和社会文化背景必须如同生物工艺学和职业训练一样加以考虑。

45．有人指出（见１９８６年5月“人口活动基金评价的比较性结果”），如果较好地理解受援国政治决策机构，许多项目本来可以获益不少。人们还发现，捐助国合作的意愿是项目获得成功的最重要的因素。但如果它们低估行政和后勤问题

的复杂性和相对地无力作出重要贡献，这就会严重影响项目产出的有效利用。人们进一步认为，应在总的经济范围内评价训练项目，因为其效益在很大程度上取决于基础设施。取决于人力利用政策是否符合国家计划和／或市场状况。

46．因此开发计划署的官员们不得不承认，如果评价符合项目逻辑和从其本身角度出发，那它必然会在范围方面受到限制。由于仅仅对按照其有益性、复杂性、类型和成本等各种标准挑选出来的一些项目进行深入评价，因此他们不得不承认，这些标准意味着评价规划采用逐个项目的办法，所进行研究的结论各有不同，结果也不易分类。因此他们不得不承认，即使对某一国家进行的一系列有代表性的评价也无法提供一种手段以确定应在方案范围内处理的总政策中的主要问题（DP/1987/21）。

47．尽管开发计划署的官员们意识到项目“持续性”的重要性，甚至有些了解被限制于其自身构架和逻辑范围内的评价所产生的相对、有限、甚至扭曲的影响，但他们似乎没有把握住其观察结果的实际意义。

48．按照联合国普遍适用的评价理论，一个管理良好的项目应能够以最低成本和最大效益尽可能迅速和正确地实现目标，使多边援助取得最佳效益并加强受益国经济发展的基础和手段。这样规定的要求是必要的，但不够充分。深入评价可以在项目完成时进行，其成本显然很高，但如果忽略或未能解决效力问题，其结果可能会无甚益处。

49．实际上，从项目的目标和具体规定来看，它可能满足关联性、效率和效力的主要标准并以最合适的方式按时得到执行，但可能会很快蜕变并失去其所有效力，因为它需要的经常开支太大或没有列入受益国预算内，因为政府的政治意图已经改变，因为社会文化环境抵销了中期效益，或者因为所采用的行政和技术手段或使用的人力资源无法保证应由受援国负责的管理方面的后续行动。另外还应该考虑到，项目的效力往往受到外汇匮乏或国际收支逆差的影响。这种意见得到了以下例子的证实。

50．一个根除蔓延杂草的项目可能执行得完美无缺，而且其成果甚至可能超出一般预料。但该项目的评价会没有任何好处，如果同时，政府的重点从维持生计转

移到商业农业部门，结果它忽视了已经建立的技术控制系统的维持，从而抵销了以前所取得的成果。

51．研究项目可能产生预期的效果。其经济和技术要素会被视为令人满意的。从关系性和效率来看，它可能是无可指摘的。但其结果仍然会受到怀疑，因为事后评价无法表明应用研究的成果如何找到市场，或如何有助于一种产品需求替代另一种产品需求。

52．训练项目可能实现其目标，完成的工作最初会产生令人满意的结果，但中期结果可能会失败。奖学金的接受者将充分得益于向他们提供的良好训练，数月之后他们将去国外谋求工作，利用训练的收获得到经济上的好处，而这种训练的原来目的是促进其本国的发展。在这种情况下，效力等于零。如果奖学金接受者即使没有涌入人才外流的潮流，却得到了晋升，结果使其训练目的所在的职位出现空缺，那么效力也会降低或偏离。

53．已经证明卓有成效的主要农业项目一旦减低其目标和受到限制，从长远来看，会对社会环境产生灾难性的影响，而且会成为政治动乱的间接原因，而政治动乱的代价与项目的预期经济效益不成比例。与此相反，同一区域另一国家的专家在他地方创造了成本极低的初级灌溉系统，任何评价都不会考虑到这种廉价的项目，但其示范性的效果却具有非常深远的意义。

54．在某些情况下，没有任何评价可以确定即使是项目的中期催化或增殖效益。讨论会或研讨会的学术水平可以由专家们正确评估。然而要确定业务活动成功与否，仅仅通过事后评价来查明受训人是否合格、专业人员是否得到重新训练以及受益国是否得到满足，这是不够的。必须查明研究小组是否已成立，技术是否已转让以及教员是否训练了需要利用这些技术的队伍。这种影响只有在任何评价进行之后相当长的一段时间内才能确定，由于这种评价往往着手得太早，因而用处很小。

55．如果比较重视数量和量度而不太重视从非常广泛的应用或缓慢的催化作用中汲取经验教训，那么评价的困难，不论其表面迹象如何，就会很严重。评价药物分发或接种运动的结果似乎比较容易。但分发率并一定意味着特定的使用率，或者根本就没有得到使用，接种的衡量由于注射次数的计算而趋于复杂，在有些情况下，

完全有可能举棋不定：进行煞费苦心的调查，这与所得结果不成比例；依靠抽样调查，则有可能导致虚假的推断。据此看来，确定矮脚稻品种产量增加量比较容易，而衡量这种增加对惯于吃非矮脚稻品种的牲畜的喂养的影响则比较困难。

影响协调和竞争

56．特定项目评价的逻辑意味着，应该更多地关心项目进程、产出和效果的特点，而不是该项目与其他项目的关系。这种逻辑促使考虑内在联系和具体效果，而不是环境和后果。即使是深入评价也只能产生有关国家经济状况的可能重要方面的不完整资料。按照评价的逻辑，必须忽视这样的事实，即把一个项目的影响同另一项目的影响分离开来是极为困难的。由于评价限于单一项目，因而无法使外部效果和影响的问题具有与后果和效果等问题同样的重要性。

57．把特定影响归因于某一项目必然纯属推测。任何项目的产出总是与另一项目相互交换和相互作用的。粮农组织旨在改善受益对象之营养的一个农业项目或教科文组织的一个扫盲项目，对于人口健康的诱导影响可能大于卫生组织单独承办或与儿童基金会合办或基础保健方案。

58．亚洲开发银行最近发表的一份报告建议，绝对不忽视需要得到资助的经济项目的人口状况和影响，无论这些项目是关于计划生育、保健、营养、农村地区人口外流、人口向中型城镇而不是向大城市的迁移。联合国每一专门机构都必须根据自身的经验处理这些问题。亚洲开发银行把这些问题相互联系起来，从而突出了各个项目在相互关系网中的交互作用的不可分离的性质，这种相互关系不能在评价的基础上得到正确的理解，因为不管评价如何彻底，总是互不相干和支离破碎的。

59．如果有人建议对联合国的补充性但形式多样的项目所推动的某一受惠国儿童健康改善状况进行回顾性评价，那么几乎无法确定哪些应归功于卫生项目，哪些应归功于饮水供应项目或采纳有利于比较合理营养的技术。

60．参加１９８６年召开的一次会议的儿童基金会的专家们得出结论说，问题不仅仅在于确定婴儿死亡率和发病率下降是否由于执行了饮水供应方案——实际上其有利影响可能扩及整个人口包括成人在内；问题在于必须承认项目对于受益对象

的直接影响是无法评估的，就项目及其效果而言，无法区别哪些应归功于受惠国政府，双边援助、非政府组织，哪些应归功于联合国执行机构。

61．对于某些机构在何种程度上通过追求一些项目中的有关但不同的目的从而促进了同一目标的实现，意见可能有分歧，这些分歧会影响到可能在不同时间进行因而从某一项目角度来看必然产生不同结论的任何评价。除了影响和把有关项目的效果详细分类的可能性之间的相互关系，双边或多边援助捐助国之间的竞争和联合国执行机构之间的对立或相互缺乏协调也会使评价愈益运作不灵和更加令人失望。

62．一个专门机构进行的一个项目，尽管从其拟订、把捐助转变为产出以及目标和预期效果之间的相互关系来看完全首尾一贯，即使与另一机构在有关领域进行的项目并不抵触，也有可能对后者起抵消作用。例如，儿童基金会的项目可能会强调保健和人道主义方面，而不太注意它在人口方面产生的影响。在同时和同地，教科文组织可能实施一项鼓励计划生育和限制生育率的教育方案。虽然这两个项目并不矛盾，但可能会减弱各自的效果，实际上尽管没有相互抵消。评价各自的作用不会有很大的意义，因为它们结合起来发挥作用的效果比每一项目的结果大得多。最初缺乏项目的协调必然会导致缺乏评价工作的协调。

国别方法的必要性

63．评价系统按照项目的逻辑发挥作用，不论逻辑的协调程度及其与方案、次级方案或一组项目结合的程度如何，因此无法摆脱部门和局部的限制。它不允许全面检查联合国在某一段期间给予某一国家的多边援助的中期效益，尽管可以从这种检查中得到的经验教训极为重要，而且只有统一的结果才能被视为具有重要意义。

64．可以想像，有些人希望遵循同质的项目逻辑和评价惯例，唯一目的是避免必需确切衡量多边援助的全面影响并可使他们免去自我批评的麻烦。他们可能担心不得不承认缺乏协调或浪费精力，而他们希望看到愈益增多的综合性配合。利用特定项目来进行评价可以说明千变万化的形势的问题，得出均衡的结论并给模糊和不确切留有很大余地，而国别方法的优点和（或）缺点是消除幻想和在继续、加强或放弃努力之前勾勒出实际情况。

65．只有在国家一级并按照所有联合国机构承办的国别办案的总数，深入评价才有可能提供受惠国以发展最需要的资料。这种评价集中注意既有时协调而又重复、偶然相互抵触的一些专门机构的并列或附加的活动而不是一个专门机构的业务活动，因此会暴露缺陷和过度情况。只有采用国别方法，深入评价（成本最高，因而应最有成效）才有价值。

66．虽然多数三方研究和深入评价是有关项目的，但开发计划署还进行其他类型的评价，有关进程、体制、主题、甚至国别方案。在上述情况下，按照其程序手册主要委托外聘顾问从事评价，以便得出适用其他国家的方案实施的结论。

67．尽管以此方式进行了努力以便超越项目逻辑，但其范围仍然有限。国别方案的评价至今主要针对的是开展中型活动和不太复杂业务的极少数最不发达国家。１９８７年尼泊尔的情况就是如此。此外，这些评价必须得到与开发计划署共同负责这一工作的有关国家政府的批准。最后，开发计划署不论在资助联合国项目方面发挥何种决定性作用，并不概括所有多边援助业务，其中有些业务属于某一执行机构的具体规划范围。

68．按照国别方法，深入评价将涉及所有联合国机构在一段期间内开展的所有多边援助业务。这种评价不会限于开发计划署处理的问题，在开发计划署负责的业务中，受惠国可能会大幅度修改其发展战略和重点，确定方案拟订中有重大缺陷，或在确保取得理想效果过程中遇到许多困难。评价应该定期进行，应有助于收集制定结帐表所需的资料。虽然，的确不是由各国来全面评价其国内执行情况，但全面评价联合国各机构提供的多边援助的多边援助是符合其职权规定的。这并不是一种新做法，因为双边援助和世界银行已经这样做了。

69．按国别方法进行评价具有双重优点，即保证更好地理解联合国项目和方案之间的相互联系，更好地了解捐助国政府和提供双边援助的非政府组织的关系以及它们之间的任何竞争。为了国际比较起见，评价可以提供有用的资料并有助于更好地理解某一区域失败和成功的原因。

70．这样拟订的评价可以获得更多的资料，方法是按照这些国家是岛国还是内陆国以及它们是属于最不发达国家还是新兴工业国家或者拥有大陆面积的国家或者

领土很小国家加以复通适用的方法，这种分类不是相互排斥的．就接受联合国大量援助而且指规数很高的拥有大陆面积的国家而言，评价可集中于某些省或区域实体。在其他国家，评价可以把国别方法同地理区域方法结合起来。

71．人们广泛认为，应以一定程度的怀疑或冷漠的态度来看待“大量的”评价。项目一旦完成，专门机构往往认为它们应从事另一项工作。如果它们没有试图或不了解如何对受惠国作必要的解释，几年后后者会产生怀疑，认为这是对其内政的不可容忍的干涉。如果有些评价的费用相当于一个主要项目，开发计划署有时也会裹足不前的。

72．实际上，联合检查组为斯里兰卡进行的唯一国别评价在为期一年的时间内，使五位检查专员和四位研究助手全时工作。以下情况也是事实：经济及社会理事会１９６９年从事了一些国别评价，世界银行同时在哥伦比亚也进行了一项评价，而开发计划署于１９７３年在尼日利亚又重复这项试验性工作，所有这些机构都认为这项工作太复杂而且成本太高，因此它们都拒绝重复这项工作。但在１９８７年，开发计划署又认为应利用国别方法对尼泊尔的项目进行评价。

73．虽然按国别方法进行影响评价的费用值得考虑，但联合国目前的财政危机不允许这种评价的论点很可能被用来反对采用这种方法的评价。应该承认，采用这种评价最初会在必须开列的预算项目下增加开支。但财政危机的起码好处是促使有关官员对既定结构和程序的边际效力和相对助益提出疑问，并鼓励或迫使他们在未改变的预算范围内开展改革，而这是以前他们从来不敢或未曾想到的。原则一旦确定，即系统必须仍然冻结，工作人员的任何调动根本不可能，而且一项开支转移给另一项开支不可想像。那么评价程序的任何调整必然代价高昂．因为这涉及经费，而无所事事却无需经费，没有会计手续来区分和计算其实际代价。

74．开展某种评价的必要性不能取决于它所带来的财政问题可以得到解决的难易程度。深入中期评价的费用已经列入其预算。自我评价的费用包括在各执行机构的一般经常开支中，而没有编制分类帐户以表明其实际开支的数额。这在一定程度上说明了为何比较频繁地使用这种方法，但还不足以证明它可成为一种惯例。与此相反，影响评价尽管很有效率，但很少予以使用．主要原因无疑是，项目文件和业

务预算不允许或不提供资助这些评价的手段。

75．尽管影响评价似乎成本最高，但可能确属最为有效。开发计划署的报告指出，评价系统允许每年决定五六个大型项目是否应结束或重建，从而证明费用不论多高都是合理的。此外，在开发计划署评价司倡导下于１９８８年２月在日内瓦召开的机构间会议上，世界银行的代表非常强调这样的事实，即没有任何发展机构有关技术合作或投资项目的影响评价。按国别方法开展的协调和总体影响评价赞成着眼于产出和发展的更为经济的项目管理。

76．仍然存在的问题是，认为按照国别方法开展的影响评价太复杂的这种论点是否能被接受。这种评价无疑是困难的，但这一事实不能用来推断它们是不必要的，尤其是因为试验这些方法的人极少而各试验之间又隔得很久，因此不容易发现有什么官员根据完整的经验对这些方法的必要性提出问题。

77．一些显然事实是不可避免的。联合国机构在执行项目时必须考虑可行性和后续行动。如果执行机构的权限范围邻接或重叠，最好确定项目之间的相互作用和综合后果，然后才查明每一项目的具体影响。几个执行机构参与同一评价过程必然会引起竞争和冲突。然而复杂因素经常出现，但决不是不可逾越的．因此不能使它成为畏惧或拒绝面对复杂局面的理由。

三．评价程序的改善

78．联合国各机构的评价单位遵循程序并采用符合指导原则的方法，而且在一定程度上隐含地参照开发计划署手册及其关于监督和评价的章节，其最近的文本是１９８７年４月发表的。本章述及而不是发动一系列改革，具有评估、准则渊源和检验正统观念的性质，代表了三年工作的成果，它所载的规定被认为在开发计划署驻地代表、执行机构和受援国政府的配合下已在外地得到检验。

手册和理论之间的时间滞差

79．已开展的系统应该既有综合性又便于管理。对结果的关心和煞费苦心地致力于建立有效的逆动和反馈机制，支配着不仅对开发计划署，而且对需要开发计划署资助和支持的各执行机构执行的项目和方案同属有效的程序。因此基本原则是通

过确保更紧密地结合评价工作所产生的资料来提高其成本效果，以便改进随后项目的拟订和执行。

80．为编写开发计划署手册新章节而进行的微妙谈判，不论其意图如何，无法给这一系统带来重大变革。经整理后，它更具优势，但程序基本上继续反映几乎机械地把投入转变为产出的概念，以及它所隐含的个别项目评价的分析性和部门性拟订。这一系统继续集中精力实现规则方面的统一，同时暂时不谈各自背景。它比较侧重合理的权限和客观性概念，而不是评价者的素质和对及时的要求。

81．尽管开发计划署手册具有某些优越性并成为一种范本，但它并没有使为执行机构制定的指导原则黯然失色，而且继续与后者共存。毫无疑问，这些文件为实现统一性而作的努力在最近几年得以保持。就基本评价词汇达成初步谅解是一项显著成就。令人鼓舞还有，所有联合国机构都已采用了联合检查组拟订并经开发计划署通过的定义。

82．但是各种手册之间仍然存在某些分歧，主要原因不是执行机构主管领域具有部门特点，而是对于评价和监督的阶段和方法的相对重要性持有不同看法。行政结构和中心评价单位采取的立场参差不齐，这也会产生影响，并往往给全貌带上成见色彩。但可以继续标准化工作，以利于确定用语和澄清概念，而程序规则的协调似乎进展迅速，因而不需要太多的注意。

83．尽管开发计划署手册的最近文本载有于１９８７年7月１日生效的程序，但目前设想审查这些程序并着手修正似乎还为期过早而且也有困难。然而令人非常遗憾的是，新的一章对于最富有创见的评价理论家们同时进行的研究未能给予足够考虑。

84．这些理论家在一定程度上对这一专题进行了新的研究，他们的结论倾向于赞成采用国别方法和更新评价工作的社会政治环境，而不是采用项目方法和把社会政治环境缩减为推断和外部因素的紧密结合。

85．事实上，这些专家倾向于把评价看成是一种社会政治过程，这种进程是与地理范围联系在一起的，而且涉及同其发展环境不可分割的地方价值观，受益对象和伙伴。他们认为，笼统地看待从这项工作中吸取的经验教训是危险的，发现差别比突出共同性质更为重要。他们认为社会文化因素与技术因素一样重要，所以希望

把它们看成是投入而不是一种令人头痛的东西。他们希望评价者不仅仅是一名技术人员、讲解员和评判者，而且还是一位谦虚耐心、思想开放和适应力强的教育工作者、讲解人、设计师和行动者。

86．这些理论家不太信任权限的准确和延伸，而比较相信评价官员和其活动正在接受评估的伙伴之间应建立的解释关系，他们十分重视合作与谈判的概念。因此人的因素以及与用户或受益人进行对话比规则和程序的不变的相对合理性更为重要。

87．尽管不象这些过分热心的理论家那样谈论第四代评价，也不认为开发计划署手册落后了一种评价，但人们认为应该从这种研究中汲取可以归为一类和可以改进目前生效的指导原则的因素。

程序的及时性

88．一项评价的可靠性应当原则上按照所采用权限范围的确切性而不是按照评价官员的个性来确定。但目前经验表明，一项评价的价值和采用的方法也会因以下因素有所不同：评价官员、其独立性、其级别和其在有关机构内外的地位及其权限的效力。

89．不用说，过度的主观性和有缺陷的方法只会削弱评价的结论。开发计划署和执行机构的手册象联合检查组以前的报告一样正确地强调了这项工作的客观性并反复提到这种必要。但所有手册似乎都没有足够注意及时性的重要，而及时性也是绝对必需的。客观性之必需是决不会过分的，及时性也是如此。

90．客观的评价首先必须在合适的时间进行，这时间因情况而异，取决于受评价项目的类型。如果要证明有用，就必须在业务活动完成之后的很短时间内进行评价，这取决于预期效果是比较精确还是比较散乱，简单还是复杂，以及它在中期或长期是否明确适合总政策。及时性还应取决于项目的主要职能性质，取决于它是否有利于创立或巩固一个机构、一项试验或试验项目、一个训练项目、紧急援助或后勤支助。

91．评价还应考虑到时间范围和项目的性质。农业的时间范围不同于工业，接种运动不能以相同于数学方案的方法加以评价。综合性农业发展项目或文化运动中

多方面和多价因素的先决条件是，影响的速度大大低于采用较合算的变种种子和在某一部门建造公共厕所。有的项目旨在全面削减国内能源消耗，对这种项目的评价时间范围不会用于有关能源节约的研究和发展项目，从长远来看，这一项目可能具有较高的成本效益和促进采用新的技术，但在短期内只对消费者及其行为产生轻微影响。

92．如果没有对深入评价规定过分严格的时间限制，看来关于开展这项工作的决定应着重考虑到可以最有效地开展工作的时期。不太早也不太晚的合适时间是捉摸不定的，没有任何手册可以表明限度和起点。但是应该在所涉经济部门和主要职能性质方面探讨评价时间差别问题。

93．开展的工作不应产生因其结论脆弱而很快过时的报告，也不应导致既矫揉造作又缺乏实际内容的脑力工作。因此开发计划署和执行机构的手册应在阐述手册使用一节，更清楚地阐明时间范围和限界，以便确保各种项目更好地利用时间。

94．评价的依据必须是不仅明确理解其目标，而还应理解执行评价的机构的性质。传递信息的技巧和工作的性质一样重要。无论评价的拟订和进行得如何出色，无论从中汲取的经验教训如何地合理和无懈可击，用户不一定能够吸取其结论或认为宜于利用这些结论。所提供的资料的有益性和所提出建议的范围将在较大程度上取决于受援国必须能够消化它们并将它们用于决策目的的社会文化环境，而较少取决于方法的性质和价值。

95．尽管权限比较确切详细，但评价干事并不总是能够掌握用户的需要。他们仅仅理解推动各机构并制约决策过程的规则，这是不够的。他们必须从用户的角度来理解—— 即使评价干事认为后者并未正确理解它们。决定是在特定社会文化环境中而不是在其他情况下作出的。必须发动的相互作用过程需要敏感性训练和群体动态方面的知识，而这恰巧是评价往往缺乏的。

96．评价不需要阐明完全的事实，也无需识破明显的谎言。它必须客观，就象必须及时一样，因此不得不避免提到某些事情。它必须衡量可传递和可忍受的事物。它必须理解权限中没有明确表明的用户的愿望，并整理其结论，以便于或相对地便于列入将要作出的决定中。

97．无论评价干事如何客观和富有经验，他必须抵制自我沉溺于智力游戏的诱惑。评价不是学术论文。用户必须在表示同意之后相信评价的效用，即使为此付出的代价是评价干事将其研究论证搞得玄乎其玄并使之完美无缺。评价报告的发表并不自动和必然地构成一定数量的知识和经验的有效转让。没有任何机制能保证一种经验教训无论如何贴切，会被理解和运用，但至少可以和应该想一想用户接受这种经验的方式和使之可以吸收和易于理解的方法。

各种评价的调整

98．如果事实上所有有关人员都认为联合国对评价予以应有的重视，那么就应该加强其基础并改进程序的协调，以便以能确保更有成效地利用成果的方式分配任务。我们不打算也无法提供详尽的解决办法，因为这需要大量专家的援助，可以通过使机构间评价问题工作组制度化为这些专家的研究创造合适的论坛。然而我们的意图是，在不超越本报告范围的条件下辨明总的方向并指明在重新安排各种类型的评价过程中应遵循的方针和办法。

99．较高层次和较低层次的、总部内外的、执行机构内部的、捐助国政府之间的和开发计划署内部的评价管理涉及如此之多的伙伴，人们理所当然地会怀疑往往针对不加区别的读者的工作最终是否有益，因为它未能适当区别用户和受援国之间、日常管理和长期政策之间以及受惠国和受益对象之间的不同。与过多的伙伴保持过多联系的愿望可能会产生的结果是，与任何伙伴也不能有效地沟通。

100．开发计划署已经意识到这一问题，既没有低估其困难，也没有对其广度缺乏认识，但现在它不回答它已经简单地提出的问题（DP/1987/27）。其更新和修订的手册具体规定了编纂和传播通过评价所获得资料的程序。但仍然有待于完成的工作是，改进方法和手段以确定和巩固据认为最具效益的成果，以求按照责任程度和受惠类型最有利地应用这些成果。

101．虽然不声称已提供解决办法，但有可能提出某些指导原则。在执行项目的受援国里，国家协调员、开发计划署驻地代表、专家和执行机构首先需要的是可以从第一批监督所获得可立即使用的最新资料。正是在这种情况下，项目实绩评价报

告和三方审查应继续成为基层核查的核心，这种核查把自我评价同监督结合起来，因而使报告结论同即将作出的决定紧密联合起来，使外勤报告的结果联系于项目执行过程中有待作出的修正。

102．在方案一级，自我评价只会有益无弊。在项目一级，其可靠性和有益性主要取决于扩大使用项目实绩评价报告，更取决于重新确定三方审查工作的方向，而根据粮农组织代表在１９８８年２月在日内瓦召开的机构间评价问题工作组会议上的建议，将来三方审查的议程应包括项目可行性和持久性问题。

103．为了避免报告、研究和自我评价之间令人遗憾的混乱，还需要一些澄清。在许多情况下，“自我评价”一词仍然用以称述纯粹是下级向上级的报告或惯常的汇报。有些不同但又是类似办法的一部分，把“自我评价”一词用于只是参与性程序的某些事物，即使采用此方法的人也如此认为。例如在难民署，这样确定的评价的责任将移交给已经与难民署达成协议的机构．“一旦有可能”就移交给难民本身。

104．深入评价的范围也应作非常明显的变动。首先，对很复杂和（或）费用超过１００万美元的项目的中期深入评价，是否有益似乎很值得怀疑。由于这些评价的经费包括在项目预算之中，而且费用不是过分高，因此可以用这些评价来中期评估项目的关联性、效率和效力。但它们忽视了可进一步证明其用途的可行性和持久性问题。如果可行性和持久性问题可以在项目拟订阶段予以解决，并在项目实绩评价报告和三方审查的范围内得到持续的审查，那么就可以免除中期评价工作，而不会招致任何重大损失。

105．事后深入评价的经费不包括在项目预算之中，这可能有一定程度的效用，因为它们提出的结论可以用于类似项目的预备或方案实施。但执行机构的经验表明，其成本很高，而它们可以得到实际使用或其预期反馈的好处并不总是能证明其为合理的。只要事后评价不是类似于简单的考古重建，它们就可有利于某种规模的方案而不是有利于项目。对项目而言，应当有选择和有限度地利用这种评价。

106．因此理想的做法是，深入评价的范围应包括已被采用的主题或体制评价。但主要焦点应是国别或地理区域影响评价，这些评价涉及项目执行和后续行动，并可解决成本效益和持久性问题。审查一组项目的成本效益和持久性实际上是让无必

要带有偏见的关联性、效力和效率的概念产生意义的最佳方法，这些概念适用于个别项目，但对这些项目的评估只会局部地或歪曲地看待联合国技术合作活动。

107．在这些条件下，按照国别方法所做评价将比按照项目逻辑方法所做评价更容易抓住要点，以便研究或审查多边援助机构和战略。同时，它将有助于将最重要的调查分析和具体建议转交受援国政府。这种评价较好、较独立和较可靠，即使它似乎缺乏外交辞令和比较挑剔，但不仅可以拟具决定，而且还为此奠定基础。最重要的是使信息适合用户，即成员国—— 其中多数是受援国。

准则的修改

108．修改而不是改写联合国程序问题手册的过程因此意味着列入某些要点和推敲其他要点，这不会影响有关评价和监督一章的编写和实质，而可使其发挥更大效用。开发计划署官员本身已提请注意的疏忽问题（DP/1987/27）也可得到弥补。

109．编入更多有助于评价干事工作的说明可改进准则。一些国家审计机构采用的审计手册概括最重大和（或）常见的不规则行为以及最常见类型的错误和疏忽。执行机构和开发计划署手册的附件可载有项目执行中遇到、在评价过程中必须评估的最明显的困难和限制的清单。

110．儿童基金会最近为其外地工作人员编写了一份清单，其中列举训练活动方面最经常发现的缺陷和弱点。可广泛采用这种设想，并可编写有关各机构具体需要的清单。清单可表明无论如何需更详细说明的职能、目标和外部因素。同样，手册应列举影响项目及其时间表实施的季节性限制（旋风、季节风等），以及有助于监督项目内容所产生成功与否的一系列进度指数。这些进度指数随着适用范围和执行机构必须从事的活动的类型而有所变化。

111．为了获得有所改进的资料并更好地掌握如被忽视可能会限制评价范围的差别和歪曲，手册还应提到为参考目的而列举的某些分析标准。例如，如果评价忽视了宗教、种族、职业或区域差别、收入水平、生境类型和自然资源，保健项目的全面效果将会微乎其微。

112．另外还应研究各种方法和手段，以期使手段更好地适合开展评价的国家的

发展性质、规模和水平。下述的情况看来是不可能的：同样的效益标准同时适用于最不发达国家和新兴工业化国家、拥有大陆面积的国家、岛国和内陆国和联合国执行机构或受援国的政府作为执行机构实施的项目。

113．最后，手册应在有关其用法说明的最初章节中强调严守时间的必要性。奢望实现完美或沉溺于周密不应妨碍评价取得成功。计划的完善和研究的拟订如果可能会延迟其筹备工作，就应予以避免。内容不太完善但按时完成的报告比内容精细但未在规定期间内完成的评价更为重要。许多报告编写的推迟影响了其理应促进的评价，这种情况突出了遵守时间范围和期限的重要性。

※ ※ ※ ※ ※