



影响评估术语表

第一版

日期：2009年12月

如对本术语表有任何意见或建议，请发送至3ie@3ieimpact.org.

推荐引用：3ie (2009) 3ie影响评估术语表。

影响评估国际动议组织：印度新德里

归因(Attribution)

在已考虑了其他因素对研究对象的影响的前提下，在结果中观察到的变化在多大程度上是干预引起的。

损耗(Attrition)

或是在干预期间有参与者从处理组中退出，或是在面板数据的收集过程中，在第一轮之后的几轮调查中未能获得其中某组数据。这两种形式的损耗都可能会使影响评估的结果产生偏误。

平均处理效应(Average treatment effect)

对受益组（或处理组）影响的平均值。见意向处理和处理后的处理。

基线调查和基线数据(Baseline survey and baseline data)

在开始干预之前收集数据的调查。基线数据对于进行双差分析是必要的，应该同时从处理组和控制组都收集数据。

前后对照(Before versus after)

见单差分分析。

受益者或者受益者群体(Beneficiary or beneficiaries)

受益者群体是指直接或间接从干预中获益的个人、公司、机构或者村落等。

偏误(Bias)

因评估方法或者抽样设计（例：不是由于抽样错误）中的问题引起的影响评估的结果与真实值相异的程度。

盲法(Blinding)

隐藏哪些对象属于*处理组*,哪些属于*对照组*的过程叫做单盲法。在采用双盲法时,不论是对象还是试验者都不知道哪些对象属于哪个组。而在采用三盲法的试验中,数据分析者也不知道对象是属于哪个组。一般而言,盲法对于社会经济发展干预是不太可行的,因此可能会产生*偏误*。

整群抽样(Cluster sample)

一种多阶段的*抽样方法*。在这一过程中,一个样本首先从某个地理区域(例如片区或者村庄)中选取,然后从选定区域中选取某些家庭、企业、机构等作为研究样本。与简单随机抽样相比,这种抽样方式将造成更大的标准误差,但是出于降低成本的考虑,这种方法经常被采用。

对照组(Comparison Group)

一组与*处理组*(或者*参与者*)的特点类似却未受到*干预*的个体。在试验环境中,评估者能够确保没有混杂因素影响对照组,那么这个组也被称为*控制组*。

置信度(Confidence level)

对影响的真实值(或者其他统计估计量)被包含在在某个特定范围内的确信程度。

混杂因素(Confounding factors)

影响相关*结果*的其他变量或者决定因素。

污染(Contamination)

指控制组的成员受到*干预*(见*溢出效应*)或者其他影响相关结果的*干预*的影响。由于大部分社区都有多个发展干预措施,污染是一个常见的问题。

控制组(Control Group)

*对照组*的一个特例,指的是评估者能够控制环境,因而能限制混杂因素。

成本效益分析(Cost-benefit analysis, CBA)

对于*干预*发生的所有成本和效益的比较,其中,所有的成本和效益均以货币来度量。成本效益分析优于成本效率分析的原因在于,成本效益分析能够分析多种结果,可以对不同领域的成本和收益进行比较(因此有助于对发展资源进行合理分配)。

成本效率(Cost-effectiveness)

对达成结果发生单位变化所需成本的分析。相比于成本效益分析,该方法的优点在于避免了对结果进行往往存在争议的价值估算。成本效率可用来比较不同项目在实现相关结果方面的效率。

反事实(Counterfactual)

在未受*干预*的情况下,*处理组*应该出现的*结果*的值。

因变量(Dependent variable)

能被另一个或者另一些变量(*自变量*)预测或者引起的变量。这一术语通常用于*回归分析*。

二分变量(Dichotomous variable)

一种只可能有两种值的变量，例如，“性别”（男性=0,女性=1）。在用于估计倾向得分的Probit参与模型中，因变量是一个二分变量，参加=1，不参加=0。

倍差法(Difference-in-difference)

见 *双差*。

双差(Double difference)

在 *处理组* 中观察到的 *结果* 的变化与 *控制组* 观察到的 *变化* 的差分。或者，相应的，处理组与控制组在研究结果之间的 *差别* 的变化。双差分能够避免由时间上不变的不可观测的因素造成的选择偏误。也叫 *倍差法*。与 *单差* 和 *三差* 相区别。

虚变量(Dummy Variables)

在 *回归分析* 中常用的二分变量。影响评估常以用虚变量表示项目参与（参与=1，不参与=0）作为自变量，以 *结果变量* 作为 *因变量*。

效应值(Effect Size)

两个变量之间的关系的大小（尤其指项目变量与结果之间）。见 *最小效应规模*。

合格人群(Eligible population)

那些符合标准，能够成为 *干预* 的 *受益者* 的人。这些群体可能是个人，机构（比如学校或者诊所），企业等等。

事前影响评估设计(Ex ante evaluation design)

在 *干预* 开始前准备的一个 *影响评估* 设计方案。考虑到对样本是否接受干预进行随机分配的可能性，以及对基期数据的收集，*事前影响评估设计* 比 *事后影响评估设计* 更具优势。也称为 *前瞻评估*。

事后影响评估设计(Ex post evaluation design)

在 *干预* 开始之后准备的一个 *影响评估* 设计方案，此时干预可能已经完成。除非是否接受干预采用了 *随机分配* 的方法，否则必须使用 *准实验设计*。

实验设计(Experimental Design)

见 *随机控制试验*。

外部有效性(External Validity)

影响评估 的结果适用于其他时间或者地点的程度。

机构调查(Facility survey)

对于 *机构样本*（通常是卫生或教育机构，但是也适用于警察局，培训机构等）的 *调查*，该类调查旨在评估提供服务所需要的所有要素的水平与质量。虽然也可以通过一个单独的机构员工调查收集关于员工的数据（例如：教师调查），但 *观察的基本单位* 是这个机构。如果某个机构调查是与家庭调查并列进行的，那么很重要的一点就是 *调查工具* 中要包含一些信息将家庭与他们利用的这个机构联系起来，从而便于数据分析。

析因设计(Factorial design)

在一项受到多种处理的*随机控制试验*中，其中一部分接受甲处理，一部分接受乙处理，一部分同时接受甲和乙处理。可能还有一部分是未受任何处理的*控制组*。

假说(Hypothesis)

关于两个变量之间关系的一种特定陈述。在*影响评估*中，假说特指*干预对结果造成的预期影响*。

影响(Impact)

干预对于受益者群体产生结果的效力。

影响评估(Impact evaluation)

对造成*干预结果变化*的*归因*的研究。影响评估有*实验性的*或者*准实验性的*设计。

影响异质性(Impact heterogeneity)

由于环境、受益者特征以及*干预实施过程*的不同造成的影响的变化。

自变量(Independent Variable)

能引起因变量变化的变量，通常用于*回归分析*中。

意向处理估计量(Intention to treat estimate)

对整个*处理组*的平均处理效应，不论这些处理组的个体是否真正接受了干预。与*处理后的处理*相对应。

内在有效性(Internal Validity)

评估设计的有效性，即，是否充分考虑并处理了诸如*样本选择*，*溢出效应*，*样本污染*和*影响异质性*等问题。

干预(Intervention)

*影响评估*的对象，包括工程、项目和政策。

逻辑模型(Logic model)

描述一个项目是如何运行的，展现从输入到活动和输出，最后到结果的因果链条。尽管逻辑模型展现了一种关于项目预期结果的理论，但并不能显示是否是该项目导致了观察到的结果。基于理论的方法是用来检验逻辑模型相关关系背后的假设。

匹配(Matching)

用来创建*控制组*的办法，根据与*干预结果*相关的特征，使其与*处理组*中的群体和个人相匹配。

荟萃分析(Meta-analysis)

对类似项目的一系列现有评估进行系统分析，其目的旨在得出一般的结论，支持某种假说，以及/或者得到项目总体效应的估计量。。

最小效应规模(Minimum effect size)

研究者认为能够用*影响评估*检测出的最小的效应规模。用来进行*评估效力计算*时以决定所需的*样本规模*。

混合法(Mixed methods)

在一个影响评估方案中同时采用定量分析和定性分析的方法。有时称为Q平方法。

N(N)

案例的数量。大写的“N”指的是总体中的案例数量。小写的“n”指的是样本中的案例数量。

结果(Outcome)

一个或者数个变量，可以衡量*干预*所产生的*影响*。

面板数据和面板调查(Panel data and panel survey)

通过连续调查而收集的数据，其中，每轮调查都是从同一组样本收集观测数据。面板数据可能会受到*损耗*，因而产生*偏误*。

参加者(Participant)

接受干预的个人、机构、企业或村落等。也称为*处理组*。

“流水线”方法(Pipeline approach)

指的是在一种*影响评估*方案中，*控制组*的人还没有接受干预，但是按计划他们也将接受干预。由于*处理组*和*控制组*都要接受干预，可以假设不存在*选择偏误*。但是*匹配*的质量必须经过检验，因为后来的*参加者*可能与之前的参加者不一样。

评估效力计算(Power calculation)

对*影响评估*所需样本规模进行的计算，这取决于*最小效应规模*和所需的*置信度*。

首要数据(Primary Data)

由调查者专门为研究项目所收集的数据。

倾向得分匹配 (Propensity Score Matching, PSM)

一种用来评估干预的影响的准实验方法。把处理组的结果和控制组的结果进行比较，而控制组是基于倾向评分进行匹配后构建出来的。倾向得分是指参与干预的可能性，它可以由Probit模型对观测到的特征进行回归得到。这些特征必须不受干预的影响。因此，倾向得分匹配法能基于多个特征进行匹配，并能将这些特征概括到一个数字（倾向评分）中。

前瞻性评估(Prospective evaluation)

见*事前影响评价设计*。

准实验设计(Quasi-Experimental Design)

用统计方法创建一个*对照组*的*影响评估*设计方法。该方法旨在确保*处理组*和*对照组*在除了干预之外的所有方面都是相同的，好比一个实验设计。见*倾向得分匹配*和*断点回归设计*。

随机分配(Random assignment)

将合格群体中的成员被随机分配到处理组或者对照组中（即，随机分配）的一种干预设计方法。也就是说，某人在处理组还是在对照组完全是偶然决定的，与他们的任何特点（观察到的或者未观察到的）毫无关系。

随机控制试验(Randomized Controlled Trial, RCT)

利用随机分配在合格群体中分配干预的一种影响力评估设计。由于参与者特征和结果之间不存在任何联系，处理组和控制组在结果上的差异可以完全归因于干预，或者说，不存在选择偏误。然而，随机控制试验也可能产生其它几类偏误，因此需要严格按照规定实施。也称为实验设计。

回归分析(Regression Analysis)

一种确定因变量与某个或者多个自变量之间的关系的统计方法。

断点回归设计(Regression discontinuity design, RDD)

处理组和控制组经确认分别位于某变量阈值两端的一种影响评估设计。这一变量可能是被项目官员用来确定合格群体的成绩或者观察到的特征（比如：年龄或者土地拥有量），或者是通过数据分析找到的区别参与者和非参与者的一个变量。断点回归设计是准实验设计的一种。

样本(sample)

研究总体的一个子集。样本是从抽样框中随机选取的。在一次随机抽样中，抽样框中的所有因子被选取的可能性是相同的，但是通常会采用更加复杂的抽样方法，需要在分析中采用样本权重。

抽样框(Sampling Frame)

研究中相关总体的完整列表。它不必是所研究国家或区域的所有人口，而是要限制在合格群体的范围内，例如：孩子在五岁以下的家庭，或者女性户主的家庭。对于机构调查而言，样本框指的是研究区域之内的所有机构。如果没有近期的样本框的名单，那么研究者就需要基于实地考察的构建一个样本框。

次要资料(Secondary Data)

为另一个目的而收集的数据，但是也许在随后的研究中可以再作分析。

选择偏误(Selection Bias)

在一个研究中，由于不同类型的人被选进处理组和控制组而产生的可能的偏误。因而，结果的差异可能是由于两组之间原本就存在的差异引起的，而不是处理本身引起的。

抽样误差(Sampling error)

当估计量是基于样本数据，而不是总体数据，得到时误差便产生了。

样本权重(Sample weights)

一种确保由样本所产生的统计量代表了样本来源的总体的方法。正常情况下都应使用样本权重，尽管对采用倾向评分匹配时如何实施仍存在争议，但它仍是一种备选权重体系。

单差分(Single difference)

指将控制组在干预之后的结果与它基期的结果进行比较（也称为*前后对照*），或指把*处理组*和*控制组*的结果进行*事后*比较。相较于双差分。

溢出效应(Spillover effects)

干预对处理组之外的单位产生的影响（不论是正面的还是负面的影响）。忽视溢出效应将导致估计结果有偏，如果存在溢出效应，那么受益人比参加者的数目要多。当溢出效应影响到*控制组*的成员时，就成为*样本污染*的一个特例。

调查(Survey)

采用（1）事先定好的抽样策略和（2）调查工具来收集信息。调查可以从个人、家庭或者其他单位，如企业或学校（见*机构调查*）收集数据。

调查工具(Survey instrument)

在调查过程中使用的一个事先设计好的表格（问卷）。一项调查通常采用多个调查工具，比如家庭调查和*机构调查*。

系统性综述(Systematic Review)

用科学方法对所有相关研究进行彻底的文献搜索，对某个特定议题（比如供水系统与公共卫生的有效性）进行综合分析，从而减少对研究方法和研究结果进行评价产生的错误。系统的综述比文献综述更加彻底，但是并不采用荟萃分析的统计方法。

处理组(Treatment group)

接受干预的一组人、企业、机构等。见：*参与者*。

处理中的处理(Treatment of the treated)

处理中的处理估计量是仅对那些确实接受了干预的人的影响（*平均处理效应*）。见*意向处理*。

三差分法(Triple difference)

对两组产生的相当的或不同的影响，即将每组与没有接受处理的对照组进行倍差得到的估计量再次进行差分。一个显著的三差分估计量揭示了*影响异质性*的存在。

分析单位(Unit of analysis)

组成总体的基本单位，这些单位被研究者选出来加以度量；它还指能够被量化的基本单位。

不可观测的特征(Unobservables)

无法被观察或度量的特征。在*准实验设计*里，不可观测的特征的存在可能造成选择*偏误*。