

教学评估方法

2013年07月06日，上海

全书脉络

- 第一章 岗位职务描述（基本方法、职务分析与培训方案设计）
- 第二章 人员素质测评（测评与培训、特征分析与培训设计、笔记测验、统计分析）
- 第三章 培训项目开发（基本方法、实施计划方法）
- 第四章 培训课程开发（开发、体系设计、教学大纲）
- 第五章 培训教材开发（开发与组织、开发与选用工作）
- 第六章 企业培训教学工作（教学管理、教学实施、培训师管理）
- 第七章 培训评估工作（方案设计、组织与实施）
- 第八章 培训质量管理体系建设（工作流程及控制、管理工作）
- 第九章 现代培训技术应用（开发多媒体教材、管理软件开发应用、仿真虚拟现实技术、远程网络建设）

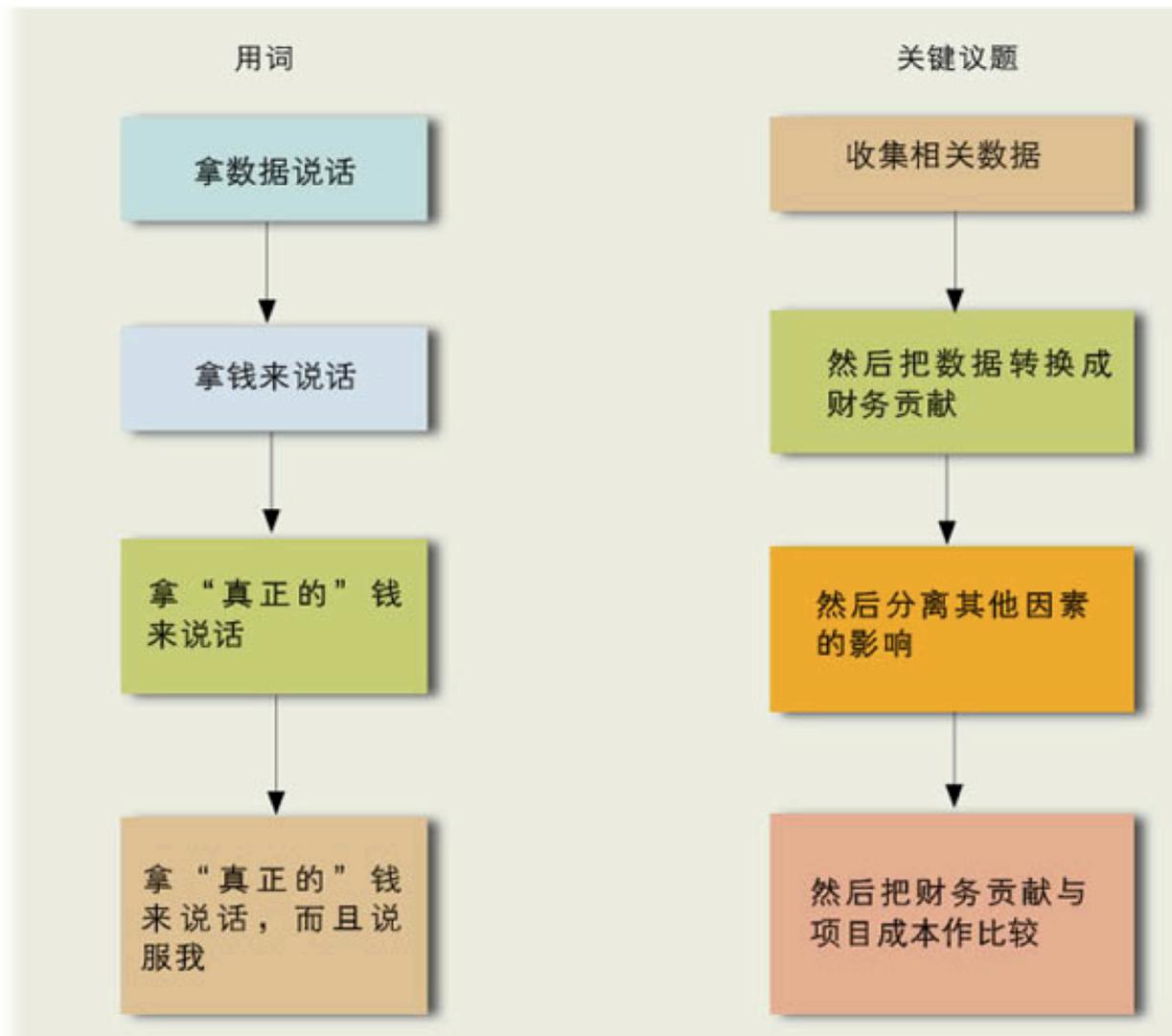
为什么做评估



培训评估的目的和作用

- 培训是否有效？
- 有效性到底有多少？
- 如何改进能更满足学员需求？





图表1 高管期望的演变

培训管理者的**发展前途**及其培训项目的**未来命运**，在很大程度上，取决于他们进行**评估的能力**以及如何使用**评估结果**！

第七章 培训评估工作

第一模块 培训评估的方案设计

能力点一：评估实施策略的制定

为什么要评估；“谁”来评估；“何时”评估

能力点二：培训评估标准的制定

确定培训标准要遵循的原则：

- 要以培训目标为原则
- 要与培训计划相匹配
- 要具体、可操作
- 要有一定的可信度

制定评估标准必须按一定的程序科学进行，不能搞“一刀切”。

- 确定评估指标。
 - 构建指标体系。
 - 设立指标等级。
 - 确定权重系数。
 - 进行试评。
-

第七章 培训评估工作

第一模块 培训评估的方案设计

能力点三：培训评估方法的选择

选择评估方案要考虑的问题：

- 评估目标与方法的匹配性。
- 评估技能。
- 评估方法和方式的选择决不能忽视调查对象。
- 组织是否认可。
- 多种方法。

能力点四：评估工具的设计

一、调查问卷的设计

（三个要点：问卷的顺序、问题的表达、问题的形式。）

二、访谈的方案设计

（访谈对象与访谈目标的确定、问题的设计、反馈式问题。）

分层次评估模式

- 反应层
- 学习层
- 行为层
- 结果层

柯克帕特里克
四层次评估模型

- 可能性和反应
- 掌握评估
- 应用评估
- 企业效益
- 社会效益

考夫曼
五层次评估模型

- 反应层
- 学习层
- 行为层
- 结果层
- 投资回报率

菲利普斯
五级投资回报率ROI模型

培训评估的种类

分阶段评估模式

加利福尼亚大学
评价研究中心
CSE评价模式

- 整个培训过程
分阶段进行评估
- 实时控制、
调整、改进培
训工作

CIRO培训效果
评估模型

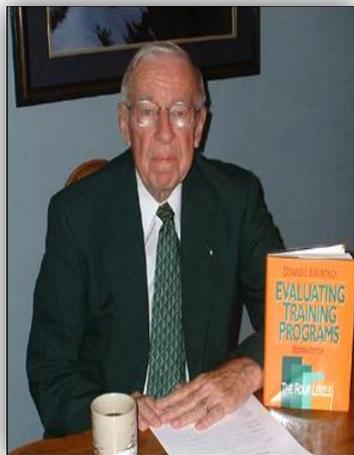
- 背景评估
- 输入评估
- 反应评估
- 输出评估

斯塔弗尔比姆
CIPP评价模式

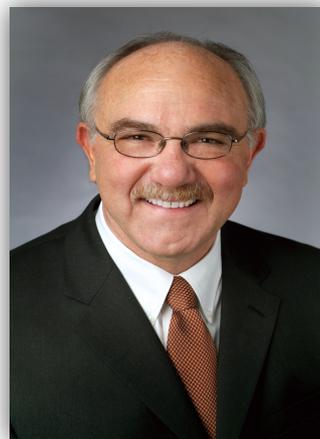
- 背景评价
- 输入评价
- 过程评价
- 结果评价

柯氏的四级评估 与 菲利普斯的五级评估

唐纳德·柯克帕特里克教授，柯氏四级培训评估模式的创始人



VS



杰克·J·菲利普斯博士，ROI研究院主席，国际绩效改进协会前任主席

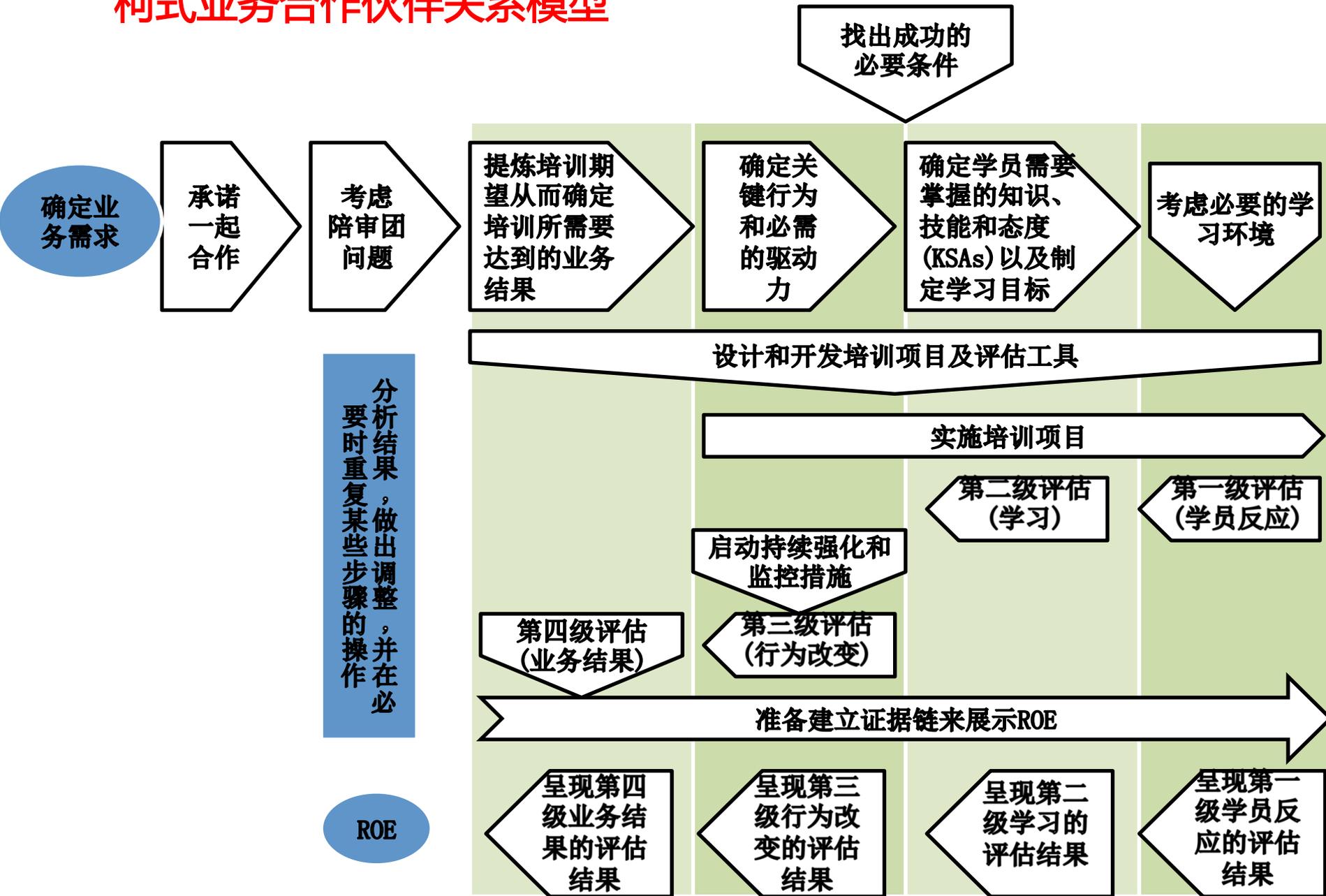
菲利普斯博士的五级评估：

- L1 : Reaction (反应)
- L2 : Learning (学习)
- L3 : Application (应用)
- L4 : Impact (影响)
- L5 : ROI (投资回报率)

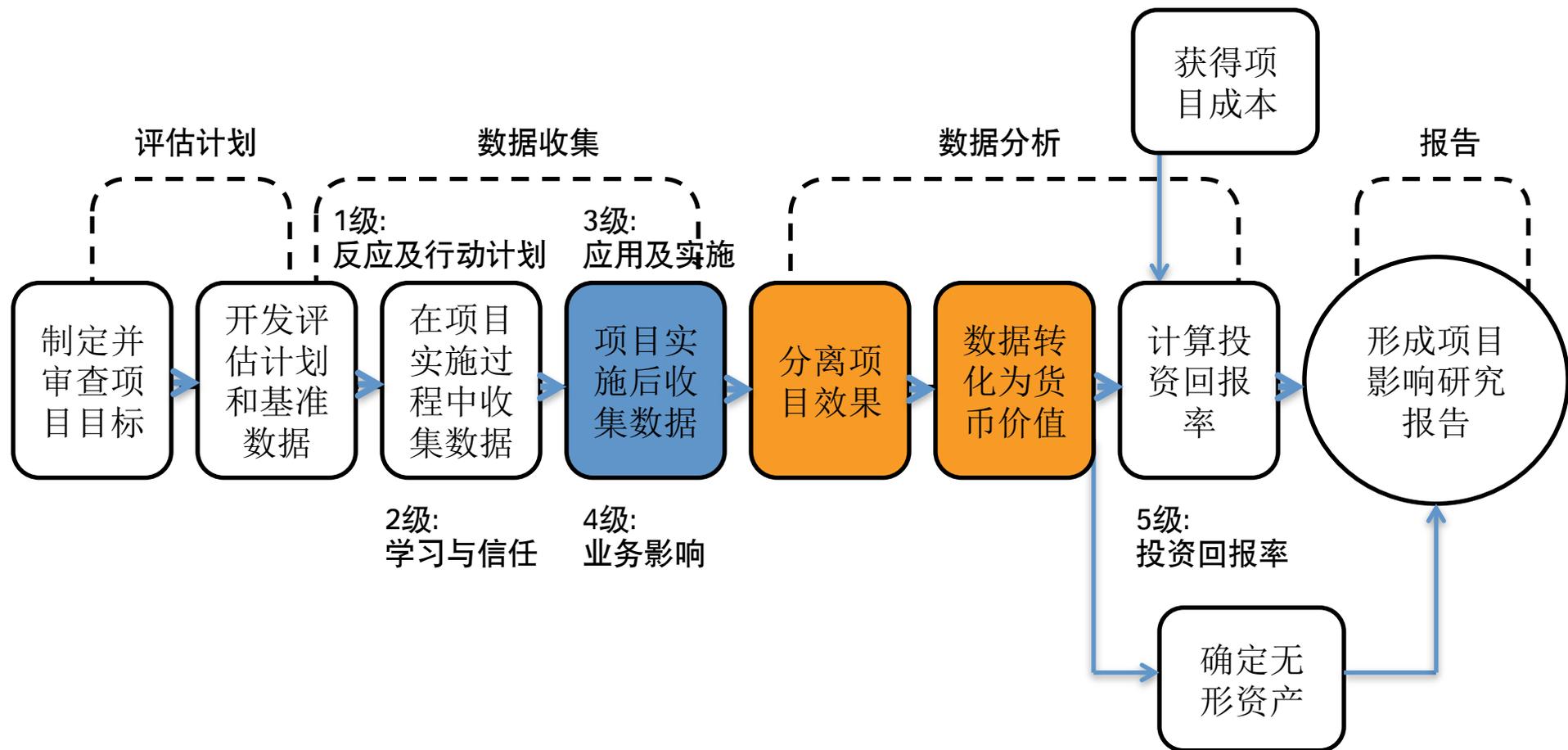


柯式	杰克
期望值回报率	投资回报率
ROE模型	ROI模型

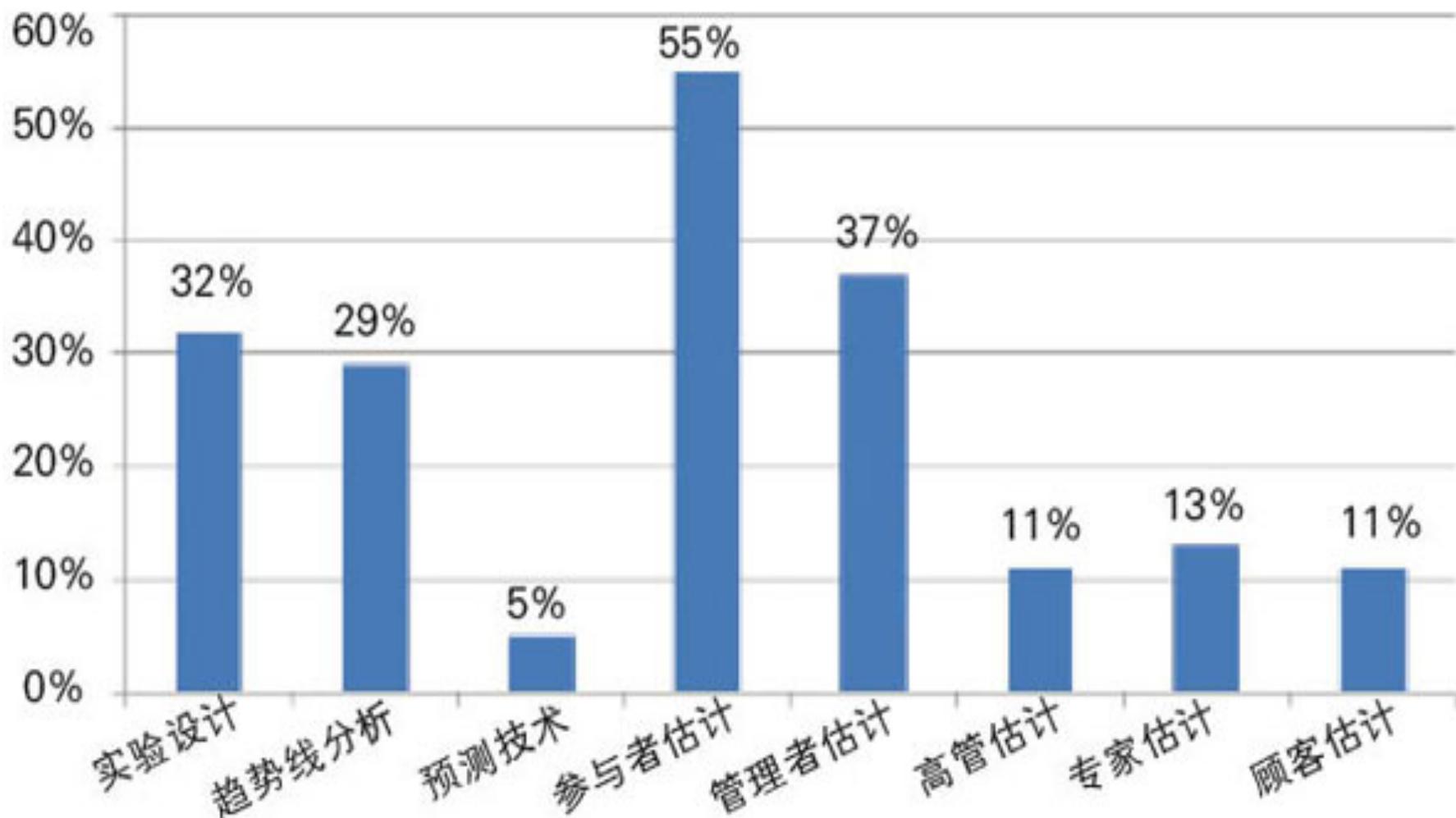
柯式业务合作伙伴关系模型



ROI 程序模型



ROI程序模型 - 贡献分离技术的实践方法



(资料来源于用户调查, 人数=235)

图表3 贡献分离技术的最佳实践

第七章 培训评估工作

第一模块 培训评估的方案设计

知识点一：培训评估标准的内涵

“以一定的方法、途径对职业培训的计划、实施以及效果等作出价值判断”时的依据。就是对培训工作质与量的具体规定，是衡量培训工作质量及其价值判断的尺度。

（诊断性、形成性、终结性三种评估方式）

知识点二：评估标准的指标体系分类

指标体系

是由系统化的、有机联系的指标构成的评估标准体系，是评估目标的具体反映。侧重于且较多地依赖于对事物的测量和计算，是用数字和量度来描述的。强调研究程序的标准化、系统化和操作化。

概括性问题描述

是以一系列较为抽象的问题作为评估标准的，类似于我们常见的调查机构。一般无法以具体的、可测的行为化指标反映，但又是重要的不可忽视的评估标准。概括性基于描述性分析，本质上是一个归纳的过程，是定性研究。它强调了研究方式和研究手段的灵活性、特殊性。

第七章 培训评估工作

第一模块 培训评估的方案设计

知识点三：培训评估指标体系的构成

素质指标（条件指标）。是从开展职业培训活动的基础和条件这一角度提出的评估指标。

职责指标（过程指标）。是从被评对象履行工作职能情况的角度提出的评估指标。

效能指标（效果指标）。是从被评对象完成任务结果的角度，从投入与产出的比例的角度确定的评估指标，它重点地考察职业培训活动所带来的质与量的变化。

知识点四：培训评估方案的设计种类

按照评估的具体实施方法。是否包括培训前后测量、是否有对照组，培训评估方案的设计可划分为不同的类型。

按照培训评估的主体划分。分为以学员为主体的，以培训单位为主体的，以及以培训组织单位、用人单位为共同主体的评估方案。

按照评估的内容和时间安排来划分。分为诊断性评估方案、形成性评估方案、终结性评估方案。

第七章 培训评估工作

第一模块 培训评估的方案设计

培训评估的实施方法简介

一、资料收集阶段。主要方法有：

1. 资料收集法
2. 观察法
3. 访谈法
4. 问卷调查法
5. 测验法
6. 总结法
7. 追踪法

二、具体评估方法有：

1. 终结性与形成性评估法
2. 纵向与横向评估法
3. 单项与全方位评估法
4. 定性与定量评估法
5. 个体评估和集体讨论评估

第七章 培训评估工作

第二模块 培训评估的组织与实施

能力点一：评估实施的一般步骤与任务

评估实施的一般步骤：一般分为：预评估、再评估

评估实施阶段的主要任务

1. 进行培训评估的动员与宣传
2. 收集信息
3. 评分（关键任务之一）
4. 汇集整理有关资料

能力点二：拟定评估方案的主要内容

- 一、明确评估对象和评估主体
- 二、明确评估目的
- 三、选择评估者
- 四、明确参与者
- 五、明确制定评估方案的框架思路

能力点三：收集相关文件与信息 认真收集相关材料、尽可能多的项目参与人交谈。

能力点四：关键要素形成共识的方法。（11点，P396~397）

第七章 培训评估工作

第二模块 培训评估的组织与实施

能力点五：评估数据的整理和分析方法

三类统计方法尤其适用，即：集中趋势分析；离散趋势分析；相关度分析。

能力点六：撰写评估报告

1. 引言； 2. 概述评估实施过程； 3. 阐明评估结果； 4. 分析与建议； 5. 附录。

能力点七：实施评估反馈的方式方法

一、充分尊重被评估者的知情权和隐私权，既要及时、准确地将评估结果反馈给当事人，又要采取适当的方式，以防评估结果不当传播给被评者为伤害。

二、为使评估，特别是对培训项目评估的价值最大化，在不涉及有关当事人隐私权的前提下，有必要在有关领导的召集下，召开一次评估工作总结会，可邀请主管领导、培训项目设计者、项目负责人、教学管理人员、培训师、参训学员代表等参加。

三、既要重视评估结果，又不能把评估结果作为衡量被评者的唯一标准，尤其不能把评估结果作为被评者业绩考核的主要依据。

知识点：评估的组织机构

一般有评估工作领导小组、评估工作专家小组、评估工作小组三类

内容大纲

- 一、教学评估导论篇
- 二、应用篇：柯式第一级反应层评估
- 三、应用篇：柯式第二级学习层评估
- 四、应用篇：柯式第三级行为层评估
- 五、应用篇：柯式第四级结果层评估
- 六、教学评估未来篇

教学评估导论篇

- ◆ 三点坚持的理由
- ◆ 一点简单的体会
- ◆ 一个评估的原则

解读柯式四层评估 - 四句话

举例：假如您是一名官员，对于刚刚提出的一条新法规，试图将驾驶的最高速度再降低一些：



1. （反应）人们会如何看待这条新法规？
2. （学习）人们能够了解法规的内容么？
3. （行为）人们会因此改变驾驶习惯么？
4. （结果）超速驾驶的人员数量会有效遏制么？

举例：针对销售人员开展新产品的培训：

1. （反应）销售人员会如何看待这期培训？
2. （学习）销售人员能掌握业务操作和营销话术么？
3. （行为）销售人员会熟练地向客户推荐新业务么？
4. （结果）客户会用么、产品推广效果好么？

练习！



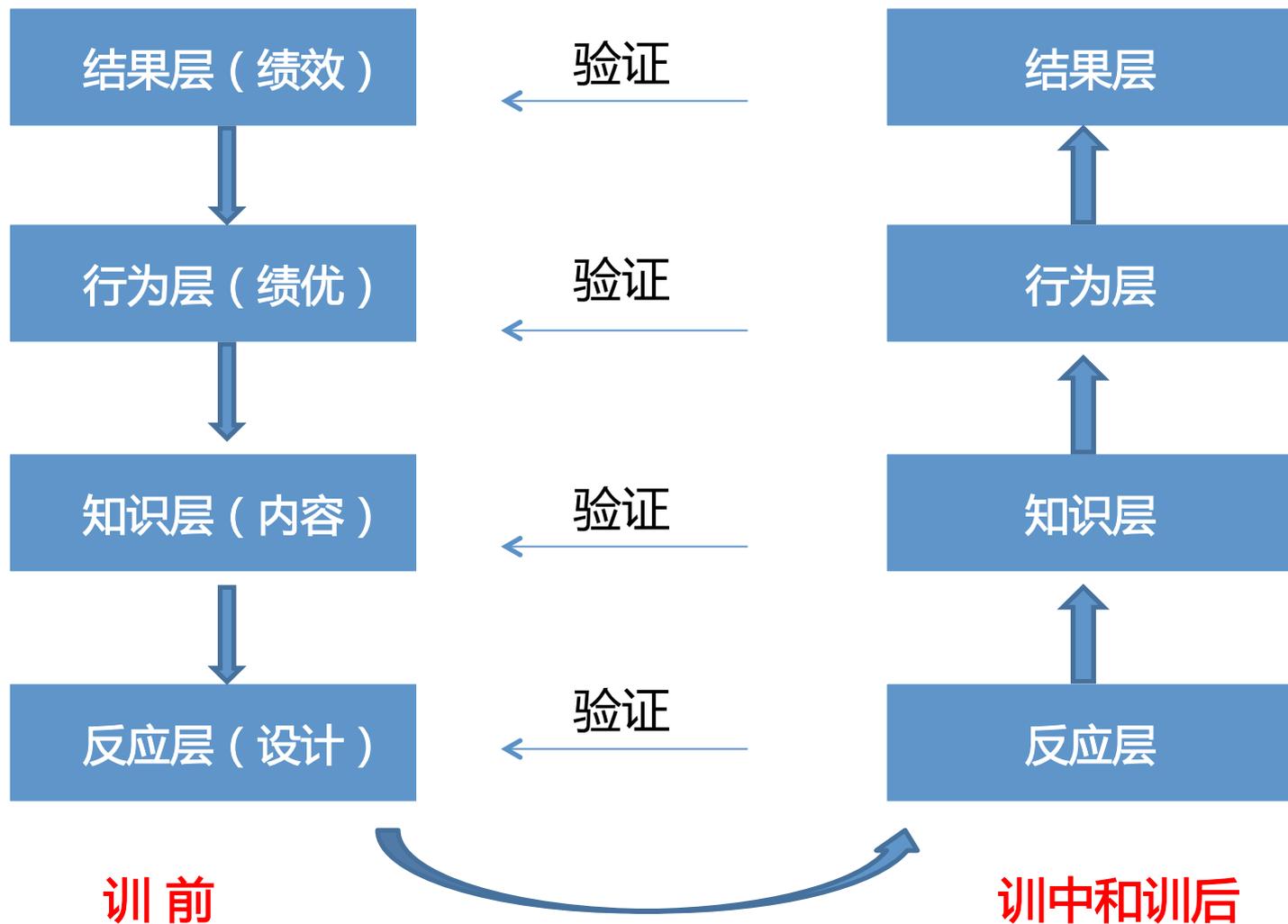
柯式评估的基本原则

基本原则：以终为始的全生命周期评估

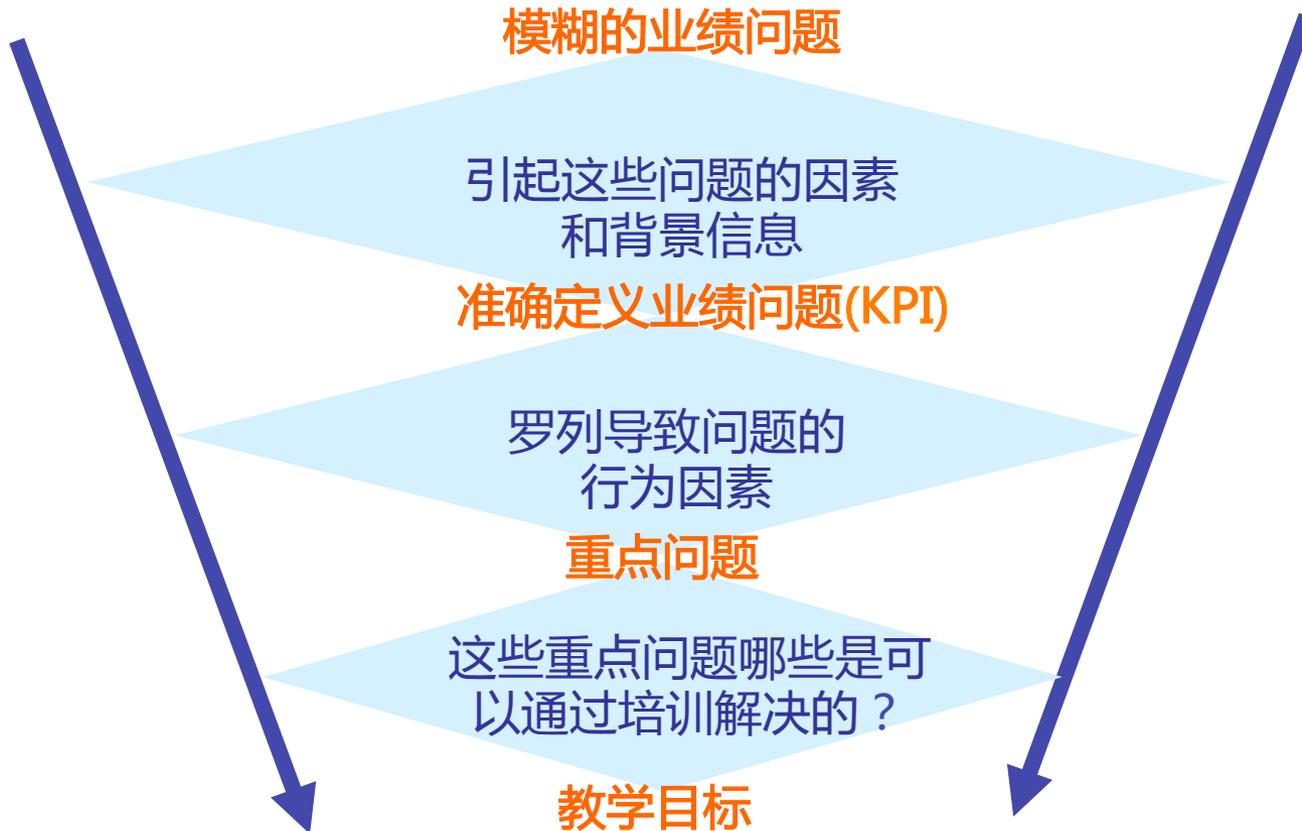
互动：分享曾经做过的评估



以终为始的全生命周期评估



培训需求分析



做对的事情 or 把事情做对？

选对

A

B



选错

C

D



做错

做对

举例：新产品培训 - 业务部门会说
我们想让所有销售人员对新产品从里到外都能有所了解。

(问) 在你看来，这个培训的结果应该这样，才算是成功的？

(答) 我想所有销售人员应当能够非常自信地将新产品的所有特征和优点告诉顾客。

(问) 如果销售人员能够做到这一点，您认为将会产生什么样的业务结果呢？

(答) 我们期望销售额能迅速增加。我们的目标是，到年底让这项新产品的销售额进入前十。



教学评估应用篇

FreeEvaluation



第一级评估：反应层评估

第一级评估：反应层评估（满意度评价）

目的和作用

- 旨在收集学员对培训项目的反应情况
- 有助于确保学员对培训活动的满意度，并且可以让学员看到将所学内容应用到实际工作中去的可能性

特点

- 第一级评估易于操作。
- 在许多组织中，100%的培训项目都要做第一级评估。

常用方法

- 调查问卷
- 访谈

第一级评估只要求学员对培训活动本身做出反应，所以仅仅根据第一级评估数据去做决定是不明智的。但是，这样的数据在确定趋势、查明项目设计中的问题、改善项目的实施效果和时机等方面很有用。

讨论和分享：你们是如何设计教学评估的？

- 哪些方面相同？
- 哪些方面不同？

在线评估



反应层评估的指导原则

确定
了解

设计
量化

鼓励
提交

即时
获得

真诚
回答

确定
标准

采取
措施

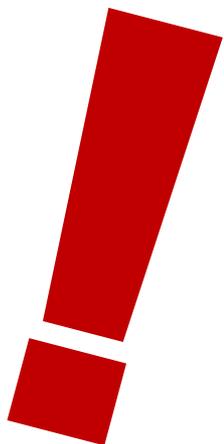
适当
沟通

创新
评估

一个发生在结业仪式上的故事

互动分享：那些年感动过我们的培训故事！





特别申明：

得到学员对培训做出的反应无疑非常重要，但更为重要的赢得学员对培训**积极、肯定的反应**，这将决定培训项目未来所得到的重视和发展程度。



第二级评估：学习层评估

第二级评估：学习层评估

目的和作用

- 为了衡量培训项目：
 - 学到了哪些知识？
 - 掌握或提升了哪些技能？
 - 哪些态度上发生了转变？

特点

- 在对大多数第二级评估流程进行开发、测试和实施的过程中都要耗费更多的时间，因此增加了整个培训项目的开发与管理的成本。

常用方法

- 测试
- 行为测试
- 模拟测试

第二级评估是不是有必要，由培训项目的性质、培训的重要性、开展第二级评估所需要的时间和资源等因素来决定。在进行第二级评估时，一定要记住一点：对知识和技能的衡量一定是以培训目标为依据的。

培训者

业绩！

管理者

成本？

消耗型衡量指标：

往往凸显了培训的成本开销



学习层评估必要么？



学习层评估的典型情景

- 培训项目所代表的关键技能或知识**必须转化到工作中去**（比如：由于法规、盈利能力或者客户服务的需要）。
- 培训项目作为获得**证书或者执照流程的一部分**，在这种情况下，知识和技能的获得是最基本的要求。
- 培训项目的资助人或者客户**要求**了解有关学习的评估结果。
- ...



学习层评估的指导原则

对照
分析

前后
评估

全员
测试

恰当
措施

课程终极目标ABCD法

常见的测试题类型

典型测试题类型	用以测试学习者在如下方面的能力
是非题	作出绝对的，非此即彼的判断。
单选题	从一系列答案中选出一个正确的答案，辨识出应属于某个范畴的选项或将某一个条目归到相应的类。
多选题	从一系列答案中选出多个正确答案。辨识适用于某对象或某概念的一系列特征。
填空题	回想名字、数字或其他具体事件。
匹配题	找出两表中所列项目之间的关联，例如事件和事件发生的原因之间或术语和它的释义之间的关系。
排序题	对一系列项目进行排序，例如按年代顺序或是按等级排序。
写作题	原创一个说明、故事、大意等。
操作题	操作程序中的某个步骤，在模拟中很常见。

主观题还是客观题？

- 主观测试题需要人工判断来评估，有时也叫做开放性问题。
- 主观题适合有教师辅导的网上课程。

- 客观题则对答案有清晰的标准。有时被设计成封闭式问题。
- 客观题通常很容易通过电脑自动打分。

对于选择题的答案

如果选项	将选项以这样的方式排列
是数字	由小到大。
是事件	按时间先后顺序。
在某一共同特性上呈现出变化	由特性的升序或降序排列。
没有内在的顺序	依字母顺序或随机排列。

为测试提供有效的反馈

1、提供完整的信息

2、温和地纠正错误答案

提供有效的反
馈

3、避免提供软弱无力的
反馈

4、在恰当的时候给出反
馈

在恰当的时候给出反馈



每道题后进行反馈

- 可以提供即时的反馈，能够提前纠正学习者的错误理解
- 可能会干扰测试的进程，对于不够耐心的学习者（尤其是那些知识水平高、很少做错题的学习者）来说，这样的反馈方式会挫伤他们的积极性



全部完成后进行反馈

- 推迟反馈到学习者已经完成了所有题目之后的方法相对更加有效，也更加经济。
- 少了些乐趣,推迟反馈也可能让一些学习者感到沮丧。对一个问题理解不当可能导致答错其他的题目



一段时间后进行人工评价

- 适用于测验题目比较复杂的情况，人工评价更有针对性。
- 不能及时得到反馈，评价推迟导致的不足。

网上教学测试的实施技巧

从题库中选择题目

- **利：**能确保学习者不会一次又一次地看到同样的试题；使作弊变得困难。
- **弊：**设计者无法以特定的顺序组织题目；要想让不同学习者和课程的测试成绩具有可比性，题库里所有试题的难度就必须相同。

试题随机化

- **利：**使得测验第二次进行时依然具有新鲜感，也让作弊变得更加困难。
- **弊：**试题之间要互不相关；无法将试题按照时间或逻辑顺序进行排列。

随机设置选项位置

- **利：**使作弊变得不那么容易。
- **弊：**要求所设计的问题不会受到选项顺序改变的影响，这时一定不要使用“以上答案全正确”这样的选项。

行为测试和模拟测试

行为测试

让学员演示在培训项目中学到的某种技能（有时会是某项知识或者某种态度）。行为测试常常用于与工作密切相关的培训项目中，学员有机会展示他们从培训中学到的东西。——例：请学员演示如何安装机顶盒。

模拟测试

与行为测试比较类似，是将某个流程的构建或者某项工作任务的应用变换成培训中所模拟的某个工作场景。这个设计出来的模拟场景尽可能接近真实的工作场景。要求学员在这个模拟的场景中完成某项模拟的工作任务、进行角色扮演练习一种新学习到的技能或者进行案例分析都是模拟测试常用的方式。——例：在讲师或其他学员配合下，请学员模拟与客户谈判的场景。



第三级评估：行为评估

第三级评估：行为层评估

目的和作用

- 测量学习者将所学知识或技能应用于工作的情况

特点

- 在培训之后收集数据。

常用方法

- 跟进问卷调查 / 在岗观察 / 跟进访谈 / 跟进座谈会 / 课外任务 / 行动计划 / 改进计划...

第三级评估非常耗时，采取怎样的方法来收集数据有赖于多种因素，包括培训项目的种类、有关人员的合作意愿、组织中存在的障碍、数据的可获得性、收集数据的成本以及数据的准备程度等。

行为层评估的必要性？



行为层评估的指导原则

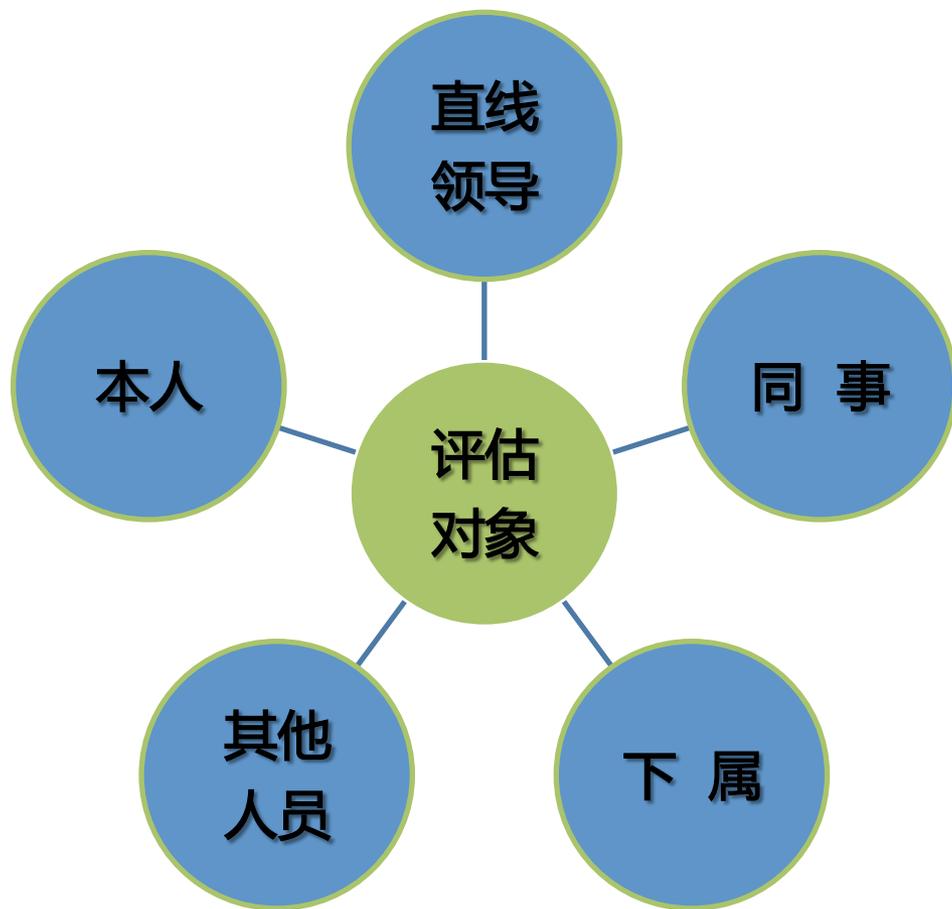
对照
分析

充足
时间

前后
评估

360测
评

360度行为评价



谁是最适合的人选？

谁是最可靠的人选？

调研问卷或访谈？

行为层评估的指导原则

对照
分析

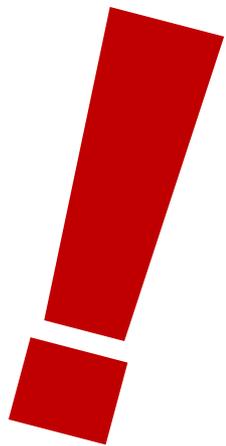
充足
时间

前后
评估

360测
评

多次
评估

成本
效益



如何促进学员将所学内容运用到工 作实践的方法： 标准行为&行动承诺

百事可乐（亚洲）领导力学院

获奖项目：储备领导人





第四级评估：结果评估

第四级评估：结果层评估

目的和作用

- 测量培训对公司或组织产生的商业影响

特点

- 非常复杂和耗时

常用方法

- 调查问卷
- 绩效合同
- 绩效监控

一般只有非常少的培训项目会进行第四级评估，而且进行第四级评估的培训项目往往是企业或组织投入很大的培训项目。

结果层评估难or不难？



难：如何跨越鸿沟？

不难：借用的衡量指标

结果层评估的指导原则

对照
分析

充足
时间

前后
评估

多次
评估

成本
效益

**优势
证据**

小结：

如何开始实施四级评估？

练习：实践柯式四层评估！

要求：

- ✓ 以终为始的全生命周期评估设计，四句话，解决方案。
- ✓ 第一层评估，细化设计，目的和改进方案。
- ✓ 第二、三、四层评估的方式和手段，达到的目的，预期的结果。
- ✓ 补充：实施的困惑和难点。

教学评估未来篇（一）



电子档案袋

教学评估未来篇（二）



互联网、移动互联网、众包模式的
教学评估实践

每日小结

1. 哪三点收获?
2. 灵感：金点子?
3. 建议和问题反馈?
4. 行动计划?



评估是为自己设定目标的过程

评估是一门科学，但更是一门艺术

评估是一个不断持续和微创新的过程