

【内部文件，严禁传播】



课程报告

课程主题：有效的教学评价

授课专家：赵德成

授课时间：2022年10月25日



目录

CONTENTS

平台介绍	/ 02
专家介绍	/ 04
现场直击	/ 05
课程金句	/ 07
课程大纲	/ 08
思维导图	/ 09
精品讲稿	/ 10



版权声明

课程报告属明德云内部学习文件

仅供VIP客户学习，请勿外传

内容仅代表专家观点

不代表明德云学堂立场



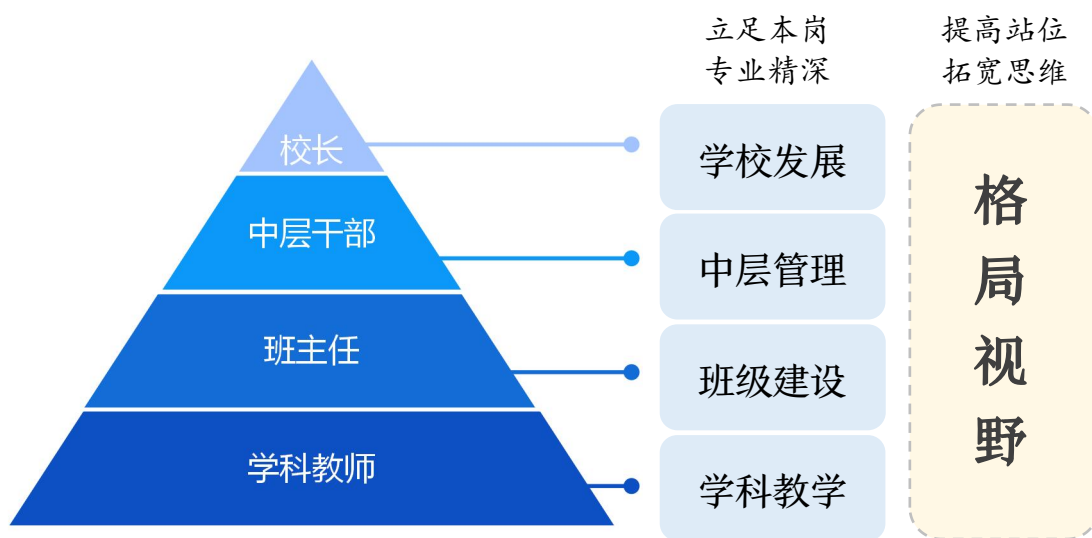
平台介绍

明德云学堂 教师专业成长智库

明德云以科技赋能教育，是教育数字化内容和平台提供商。

明德云学堂是明德云旗下专注教师专业发展的在线教育智库，聚焦前沿教育理念，秉承“助力教育数字化转型”的品牌使命，通过系统的培训体系、高端的培训内容、精细的培训服务与在线学习平台有机结合，推出“个性选课，集中学习；线上内容，线下服务”的特色师训学堂，以面向未来的培训理念和方式，致力于打造优质前沿的线上教师培训平台与教师终身学习平台。

· 课程架构



双线发展：“专业精深”与“格局视野”双线学习路径，培养复合型教师人才

分层分类：构建分层分类系统化培养体系，加强专业纵深，助力全员持续成长

· 课程内容

明德云学堂抛弃浮躁的碎片式培训和短期式培训，坚持系统性、持续性的课程编排，以“线上直播，名家引领，分层学习，集中研讨”的教学模式，提供50节高水平优质课程。

主题一 格局 视野	1	最新国际形势
	2	当前经济形势
	3	如何提高心理免疫力
	4	2022 全国两会看教育
	5	中共二十大精神解读
	6	教育数字化转型
	7	传统文化传承
	8	双减政策落地
	9	五育并举（体育）
	10	五育并举（美育）
	11	五育并举（劳动）
	12	芬兰的现象式学习

主题二 办学 治校	1	学校规划管理与校长领导力
	2	课程与教学领导力
	3	学校德育管理策略
	4	自媒体时代舆情应对
	5	学校法律风险防范
	6	学校校园安全管理
	7	教育评价体系构建
	8	学校文化建设
	9	引领教师成长
	10	名校建设经验
	11	中层干部执行力
	12	中小学党建

主题三 德育与 班级管理	1	师德师风
	2	班级管理
	3	班级活动
	4	学生德育
	5	问题孩子教育
	6	打造魅力班会课
	7	积极心理教育
	8	名班主任实践
	9	家校共育

主题四 教学与 课程 建设	1	项目式学习（PBL）
	2	教师微型课题研究
	3	有效教学与有效学习
	4	校本研修
	5	课堂改革
	6	合作学习
	7	深度学习
	8	思维育人
	9	教育科研

主题五 新教师 培训	1	新教师基本功
	2	教师专业成长

主题六 教师职 称评定	1	论文写作与发表
	2	如何开展课题研究

主题七 学生 教育	1	励志教育
	2	生命教育

主题八 家长 教育	1	“双减”后如何做父母
	2	亲子沟通技巧

专家介绍

赵德成

- ◇ 北京师范大学教育学部教授、博士生导师
- ◇ 主要从事学习评估、教育测量、教师专业发展、教育人力资源管理等领域研究。近年来主持教育部、北京市教科规划办、北京市教育学会等单位委托课题多项，在 SSCI 及 CSSCI 期刊上发表中英文学术论文近百篇，出版中英文著作 5 部。
- ◇ 2015 年年度中国人文社科最具影响力青年学者；2017 年与 2019 年被评为中国哲学社会科学最有影响力学者，2016 与 2019 年人大复印报刊资料重要转载来源作者。

有效的教学评价

赵德成

北京师范大学教育学部教授、博士生导师
曾被评为中国哲学社会科学最有影响力学者

年度课程

大咖零距离 尽在明德云



现场直击





课程金句


赵德成《有效的教学评价》
课程金句


一个好的教学评价
不是面面俱到的评
价。


赵德成《有效的教学评价》
课程金句

评价学生也要找
“KPI”，抓重要指
标。


赵德成《有效的教学评价》
课程金句

我们在评价中要看
到智之外的德、体、
美、劳。


赵德成《有效的教学评价》
课程金句

孩子们的获得感才
是评价教学好坏最
重要的标准。


赵德成《有效的教学评价》
课程金句

让学习真正发生。


赵德成《有效的教学评价》
课程金句

不以结果为目的的
注重过程都是耍流
氓。


赵德成《有效的教学评价》
课程金句

如果我们的目标是
模糊的,那么和没有
目标没什么两样。


赵德成《有效的教学评价》
课程金句

从“走近”生产生活
到“走进”生产生
活。


赵德成《有效的教学评价》
课程金句

让学生成为学习的
主人,以学定教、因
材施教。

课程大纲

一、导入部分

（一）新时代评价改革基本原则

1. 改进结果评价
2. 强化过程评价
3. 探索增值评价
4. 健全综合评价

二、对教学的终结性评价：考试

（一）现代教育评价定义

（二）泰勒模型

1. 确立宽泛的目标
2. 对目标分类或排序
3. 以操作性的语言界定目标
4. 寻找证明目标实现的情景
5. 开发或选择测量与评价方法
6. 收集数据资料
7. 把数据资料与既定目标对比

（三）加强教学评价研究

三、对教学的过程性评价：听评课

（一）把课堂还给学生

（二）学会无条件地鼓励

（三）教学评价的焦点：学生

（四）反思教学：以课为例、以课为镜

思维导图



精品讲稿

赵德成：有效的教学评价

各位老师，大家好，今天我们一起探讨教学评价的问题。

教书育人是老师的天职，每一个老师都希望上好课、教好书、育好人。那什么样的课是好课？什么样的教学是好的教学？怎样开展教学评价？这些都值得研究。这些问题研究得好，就可以发挥教学评价的导向、诊断和发展的作用；而如果研究得不好，不仅不能帮我们诊断、发现教学中的问题，而且可能会误导我们的教育。

那么该如何进行教学评价呢？我们可以从2020年中共中央、国务院颁布的《深化新时代教育评价改革总体方案》文件中找到思路。这个文件对于未来的教育评价改革基本原则做了阐述，其中提到，我们要坚持科学有效，改进结果评价，强化过程评价，探索增值评价，健全综合评价，充分利用信息技术，提高教育评价的科学性、专业性与客观性。

从这段话可以看出，未来我们需要注重四个评价，分别是结果评价、过程评价、增值评价以及综合评价。而实际上我们在推动这项工作的过程中，还要注意一个原则，就是要事优先。因为四个评价有点多，我们需要找重点。

通过找重点，我们会发现增值评价前用的动词是“探索”，也就是说增值评价还不是一个非常成熟的技术，在其推广的过程中还面临一些挑战；而综合评价前用的动词是“健全”，这意味着我们已经开展综合评价很多年了，有了一定的经验，未来只需要在此基础上进一步健全。从这个角度来看，增值评价和综合评价可以作为次重点。那么重点就剩两个了，一个是结果评价，一个是过程评价。如果我们在结果评价和过程评价之间再找一个重点，是结果评价更重要，还是过程评价更重要？或者我们把这个问题拉得近一点——对于教学评价来说，结果评价更重要，还是过程评价更重要？

我知道很多人会举手说，过程评价更重要。我曾经也问过很多校长和老师这个问题，他们多数人都说过程评价比结果评价更重要。

几年前，有一个高端幼教集团新接手了一个高端设计幼儿园，它的基础设施非常好，集团想把这个幼儿园打造成全国有影响力的优质民办幼儿园。集团的第一件事是想招一个幼儿园园长，我们被邀请去当考官，协助招聘。在招聘时我们出了一道题，要当幼儿园园长就意味着要评价老师的工作，而评价老师的工作既要全面，也不能没有重点。那么如果抓重点，也就是抓 KPI 指标，需要抓哪三个？对于这个问题，很多老师可能会从德能勤绩大框架下找三个指标，比如师德、爱心、责任心、奉献精神、团队意识、教学能力、教学基本功、科研能力、家校沟通能力等等。

那么最后我们招聘的园长是怎么说的呢？他说，如果他来做园长，他会找三个 KPI 指标来评价老师的工作，第一个是幼儿保留率，第二个是幼儿出勤率，第三个是家长满意度。大家不难发现，我们刚刚提出的爱心、责任心、团队精神、科研能力操作性不够，难以实施评价，而这位园长所说的三个指标就有操作性了。

这位园长解释，一个老师教得好不好，最终要看结果怎么样。比如，期末时要评选一些优秀教师，优秀教师在未来职称晋升上有优先权，年度的奖金额度也更大。于是，会有老师说自己早出晚归，披星戴月地工作，非常敬业。而且基本功好，获奖众多，发表了论文，有几节课也获评优质课，所以今年应该被评为优秀教师。但是，教师评价既要全面，同时也要看 KPI。

第一，幼儿保留率。一般来说，幼儿园在开学的时候，每个班招的学生有三十人左右。假如教了一年，三十个孩子还剩十八个，把孩子都教没了，这类老师评优秀合适吗？假如隔壁班的老师，基本功一般，没有参加大赛获得奖项，但是他们班的三十个孩子一个都没走，全部留下了。在这样的情况下，除非第一位老师有证据表明，流失的十来个孩子跟本人教学没关系，而是由于周边企业整体搬迁等其他原因，那么获评优秀教师才说得通。

第二，幼儿出勤率。幼儿出勤率很重要，幼儿不出勤，保育费就不能足额征收。幼儿园按天收费，如果一个孩子一个月才来十天，那么保育费必然不能足额

征收。倘若早餐费、午餐费、晚餐费都不能足额征收了，幼儿园要怎样维持运转呢？当然，排除钱的因素，如果幼儿不出席，那么我们是不是就没有机会把他们培养成社会主义建设者和接班人了？所以老师得保证出勤率，如果出勤率低，一定程度上也反映了教学吸引力、教学成效不行。

第三，家长满意度。我们要办老百姓满意的教育，如果家长不满意，那肯定不行。然而有些老师，甚至有些校长反对用家长满意度来评价老师的教学工作和学校的教育质量。事实上，这类状况有点类似于银行的顾客满意度评估项目的，银行经理、银行职员也强烈反对让顾客进行工作满意度评价。顾客既不了解他们的工作，又不是专业人员。银行工作有很多的程序，需要很长的时间，倘若顾客觉得等待时间长而不满意，就给差评，这合适吗？可是，一旦银行开设顾客满意度评估项目，银行的服务速度、服务质量真的会大幅度地明显提升。

那么，学校中关于教师评价和学校评价也是一个道理。我们应该重视家长的意见，家长也会比较客观地从总体上对老师的工作进行评价。倘若把这三个指标综合起来，对教师的工作进行评价，我相信园长能够带出一支有战斗力的队伍，能够办成一所真正令百姓满意的幼儿园。

在我看来，评价老师工作的时候，其实还是要注重结果的。如果不注重结果，没有结果导向，没有目标导向，那么工作产出的最终结果可能没办法真正做到令百姓满意。

有一天我到一所初中听课，校长问我对课堂的感受，我说挺棒的，老师基本都挺好。校长说，这所学校的年轻老师，干劲足，有朝气，有活力，集体备课熬到十点钟。他们的课堂能结合新课程理念，有三维目标，有核心素养，有教学重难点突出，不仅问题设计得循序渐进，也有小组探究活动。可是，一旦涉及做题，学生还是不会。

有多个学校的领导和老师表达过类似的烦恼，可是您听这烦恼是不是有点别扭？他说他们老师教得可好了，但是孩子不会做题。可是连做题都没教会，也就是说，最基本的知识技能都没让孩子掌握，更别说现在强调的核心素养、立德树人、全面发展了。所以，教得好不好，归根结底还是要看结果。上了一个单元或

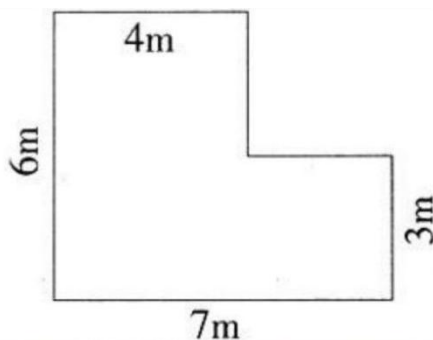
者上了一节课,孩子学会知识技能了吗?他应该具备的能力或者核心素养提高了吗?孩子的情感、态度、价值观是不是发生了我们所期待的变化?或者用现在时髦的话来说,学习真正发生了吗?这才是我们判断教师教学好坏的最重要的标准,所以结果评价很重要。

各位老师有机会不妨再去看看《深化新时代教育评价改革总体方案》,文件里强调的原则是改进结果评价,强化过程评价。也就是说在四个评价中,它首先谈的是结果评价。

政策是高度凝练的,重要的事一定放在前面讲,由此可见结果评价是重要性。可如果我们像过去那样,唯分数论、唯升学率、唯论文,这是不合适的。现在我们必须用立德树人、五育并举的教育观、全面发展的质量观来引导结果评价,让学生在知识技能以及更丰富的领域里成长 and 变化。

终结性的结果教学评价主要是通过测验、考试来进行的,而测验考试如何改进呢?我们先看一道题。

小学五年级老师一般在教完长方形、正方形、平行四边形、梯形等规则图形面积的计算方法之后,需要讲授用割补法计算组合图形、不规则图形的面积。有些老师会找教材里的例题来评价孩子的学习成效。比如:智慧老人要装修自家客厅,他家客厅有多大面积?



有些老师认为这个题好。第一点,能反映孩子对于知识技能目标的掌握程度。第二点,结合了生活,能反映孩子在解决能力这一核心素养上的表现程度。

可是这道题真的这么好吗?首先这道题能反映小学五年级孩子学完组合图形面积之后知识的掌握程度吗?如果把这道题给小学三年级的孩子来做,他们能不能做?三年级孩子只要学过长方形和正方形的面积计算方法,他们做这个题没

有任何问题。一个五年级老师在教完组合图形面积这一课之后，拿这道题考自己班的孩子，以此来证明自己教得多好，会不会有点自欺欺人？

其次，这道题之所以能反映孩子的问题解决能力，是因为它结合生活了。可是题目中提到的智慧老人都不是具体的人，那怎么算结合生活？事实上，过去的教学在结合生活的时候，只是“穿衣戴帽”而已，结合的都是虚构的生活，距离真实的生活有点远。当然我们在设置题目的时候，不可能100%真实，但是要尽量模拟得靠谱一点。

那如果让我来改进，我会把题目改成这样：今天晚上回家之后，请你量量家里每面墙的长度，粗略绘制家里房屋的平面结构图，并计算房屋面积，并于下次课在全班举行分享活动。

现在“双减”要求减少家庭书面作业的数量，所以这种实践类的作业是被提倡的。但有的老师会认为题目不好，如果学生家没有那么长的尺子怎么办？够不着房顶怎么办？其实粗略测量房屋面积的方法有很多，如果一个五年级孩子连这个问题都解决不了，必须要家长帮助，那么我觉得我们的教育真的该反思了。

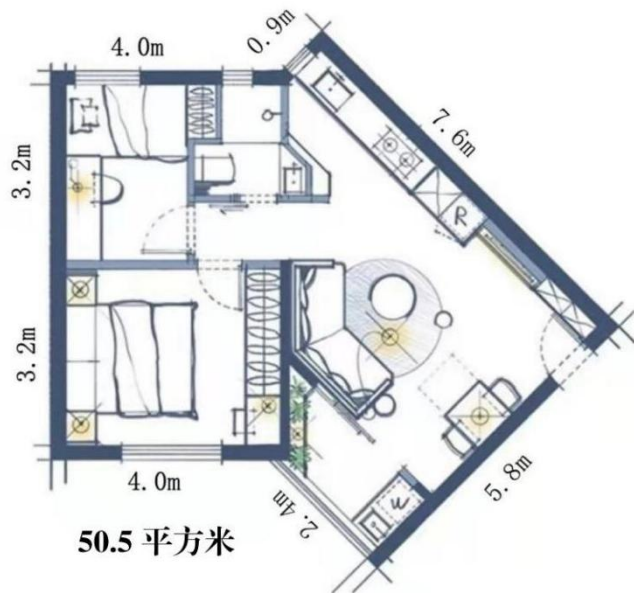
从这个角度上来说，这道题似乎还可以。有的老师认为不仅让孩子动手了，而且用数学解决问题，聚焦了核心素养。可是我想说，我认为这个题很糟糕。

理由是，在全班进行分享的时候，班里有孩子举手举得很高，因为他们家房子大，全班同学都投以艳羡的目光。可第二天，当地纪委进了他们家，查他的爸爸为什么住那么大的房子。又或者，班上有的孩子家里房子小，甚至没房子，只租了小单间，没有组合图形，面积也就“ $3 \times 4 = 12$ ”而已。他听说别人家的房子几百平方米，是不是就会着急，觉得全班同学会瞧不起自己。这道题让孩子左右为难，产生伦理的问题。再深究，还会涉及隐私权的保护问题。

这样想来，其实老师经常侵犯学生的家庭隐私权？比如语文老师布置孩子写作文，以“我的妈妈”为题，要求写真人真事，有真情实感。但是有的孩子有妈妈，有的孩子没有妈妈。离婚家庭、留守家庭，各种情况都有。老师让学生说真人真事，可是他不想让别人知道他们家里的真人真事，这是不是侵犯家庭隐私权？因此描写人物的作文完全可以换个人物，而不必局限于家庭。

到了五年级，英语老师让同学们准备一张家庭合影，并以“my family”为主题，用英语进行介绍。比如“这是我的家庭合影，我家有三口人，这个是我妈妈，我妈妈是名老师，她喜欢唱歌，我爸爸……”。可是您想想看，他妈妈是老师还是蔬果商，是地地道道的农民或者是什么职业，这是不是人家的隐私？再者说，有的孩子父母离婚，他们家根本就没有家庭合影，这又该如何？

因此，考虑到多重因素，我觉得刚才那道题目可以这样修改：张爷爷买了个小单元房，房产证上写着这套房子的面积是 63 平方米，可张爷爷感觉家里没那么大，他就量了量家里几面墙的长度（如下图），请问开发商是不是量错了？



这道题有几个特点。其一，这个题目说的是张爷爷，而不是智慧老人，不是动画故事里虚构的人物。其二，张爷爷买了个小单元房，而不是买了大别墅，没有不良价值观导向。其三，张爷爷的户型图是真实的，而不是随便画的组合图形。其四，这道题对于五年级孩子具有适度的挑战性，可以将组合图形分成若干长方形，及若干小三角形，也可以将右部直接视为一个梯形。

这道题目具有情境性、应用性、探究性和开放性，与当前中高考命题改革的趋势相对应，也体现了教育评价改革的基本特征。

之后，孩子们计算发现，该房屋面积为 50.5 平方米，那么张爷爷的房子量错了吗？要是量错了，差了十来平方米，那不是小事，需要去找开发商。

可是再想想，张爷爷家房产证上写的是建筑面积，可建筑面积还包括公摊面

积,所以张爷爷量的不是建筑面积。那么张爷爷量的是不是套内面积呢?也不是,套内面积包含了家里的隔墙、屋外的外墙等。所以张爷爷量的既不是建筑面积,也不是套内面积,而是老百姓俗称的使用面积。用使用面积除以建筑面积进行粗略计算,就大致可以得到这个房子的得房率。得房率超过零点八,没有违反政策,所以不能说量错了。

在我看来,这类题目才是结合生活生产实际,让孩子将学科知识应用于解决生活中的问题。过去我们只是“走近”生产生活,“近”是“接近”的“近”。而现在我们是真的“走进”,“进入”的“进”。

与刚才那道题配套,又有这样一道题:张爷爷发现房子买得值,于是想动手装修,给房子通铺地砖。地砖现在流行通铺,中间没有接口,这种地板砖边长六十厘米,请问张爷爷要到建材城买多少块砖?请你提出建议并说明理由。

有人会粗略计算,50.5平的房子,除以每块砖的面积,得出砖数是140.277块,有循环小数,然后把小数部分舍掉。这时,老师会说学习生活中的数学,不能见到小数就四舍五入,如果把小数舍掉,那张爷爷家的砖不够铺了。在这个时候,小数部分虽然没有达到五,但是也要“入”,所以要让张爷爷买141块砖,这才叫学习有用的数学。

可是只买一百四十一块砖够不够?可能还不够。因为砖一切割就有废料了,更不用说张爷爷家的工人从楼下往楼上搬的时候,不小心碎了几块。因此装修工人一般会提醒,要往上多买几块,而多买3%,还是5%,还是10%,要根据家里的房屋总面积、大小、形状等因素综合确定。如果让张爷爷买二百块,这太浪费了;让张爷爷买三十、五十块,肯定也不行。所以得有一个合理的区间,并且还要对所选区间进行说明。

试想,如果平时教数学、教物理、教化学、教地理,或者教语文,各个学科都有情境、有应用、有探究,是不是孩子就爱学了,是不是就具备我们要求的学科核心素养了?

所以改进结果评价太重要了。当然,要想改好结果评价,我们还要深入理解教育评价,掌握基本的教育评价模式。

接下来我给大家介绍一个全球最有影响力的教育评价模式，该模式由泰勒提出来的。泰勒说，**教育评价的本质**就是确定课程与教学大纲在实际上实现教育目标程度的过程。比如，教完一个学期后，学生有没有把该学的学会，或者教完一个单元后，评价一下这节课，这都叫教育评价。

这种评价要想做好，可以用**行为目标评价模式**。因为泰勒开发的这个模式很有影响力，所以后人就直接把它称之为“泰勒模式”。

泰勒模式有七个步骤，分别是确立宽泛的目标、对目标进行分类、用操作化语言界定目标、寻找证明目标实现的情景、开发选择测评方法、收集数据资料以及把数据资料与既定目标做对比。逐个来看，我们会发现前三个目标由粗到细，由概括到具体，逐步明确教育目标。而且大家注意，第三条说用操作化语言界定目标，也就是说，目标的表述必须“smart”。“smart”是五个单词的首字母缩写，“S”是“specific”，即明确的；“M”是“measurable”，可测量的；“A”是“achievable”，可达成的；“R”是“realistic”，即具备现实条件的；“T”是指“time bound”，有时间界限的。也就是说，目标必须是在某个时间范围内达成的，是具体可观察、可操作、可测评的。这样看来，有些老师的目标没有做到这一点。

比如，一位高中英语老师讲“New Australians”，也就是新澳大利亚人。这堂课的目标要求掌握有关单词、理解课程内容、培养合作精神。这位老师设置了三维目标，可是这些目标不够操作化，不够“smart”。课堂目标要求掌握有关单词，那么哪些单词是有关单词？这些单词要掌握到什么程度，是听说读写都会，还是只需要明白意思？如果单词有五种意思，五种用法，是需要知道五种，还是知道一种就可以？这都需要操作化界定。接下来再来看理解课文内容的目标，什么是理解课文内容？这位老师的意思是要读懂，那怎么判断学生有没有读懂？最后再看培养合作精神，是不是更加没有操作化？所以这位老师的目标是模糊的。

泰勒的研究生布卢姆认为，如果我们的目标是模糊的，那就和没目标没什么两样，它只是一个美丽的希望而已。所以我们要将目标操作化。当进行了一段时间的教学，需要检验教学成效的时候，我们就需要去设置情境和方法，收集分析

信息，然后再把这些信息和既定目标做对比，以此判定目标的达成程度。这就叫做行为目标评价模式，就是泰勒模式的大致操作内容。这个模式非常重要，无论是普通老师，还是教研员，都必须深入掌握这个模式。

举个例子，假如现在各位是中学语文老师，我们要评价学生的口语交际能力。口语交际能力非常重要，它是语言建构与应用核心素养的一部分。如果我们用泰勒模式评价学生的口语交际能力，首先要确立宽泛的目标，对目标进行分类排序。对于这两个步骤，课标要求的是，学生需要具有日常口语交际能力，学会文明得体的交流，发展合作精神，大方向没问题。

接下来要用操作化语言界定目标。这节课的操作化目标是能注意对象场合，学习文明得体的交流，耐心专注地倾听，根据对方的话语、表情、手势来理解对方的讲话意图，在交流过程中能按需调整表达方式，不断提高应对能力。这个教学目标还是比较具有操作性的。

老师带着目标教了两个礼拜后，如何判断教得怎么样？学生学得怎么样？就要做结果评价，需要设置情境和方法。所以有位老师从教材里找了这道题——每个人都有自己心里的秘密，今天晚上回家后，请你和你的父母说说你心里的秘密，你会如何说呢？请你把它写下来。

各位，你们觉得这个题怎么样？有人说挺好的，让一个初中生跟他父母说说心里的秘密，可以加强亲子沟通。它能加强促进亲子沟通吗？哪个中学生回家会愿意跟父母说心里的秘密？即便说了，也真的促进了，但是我们出这道题的目的是进行教学评价，要评价一个老师教学后让学生达成教育目标的程度。根据刚才列出来的目标，我们的目的不是促进亲子交流，所以这样做就跑题了。我们再围绕目标来看，这个题目好不好呢？让孩子说秘密是不是有点不合适？因为那是孩子的隐私，怎么能直接让他说秘密呢？往大了说，这叫侵犯学生的隐私权。

又或者，如果他是留守儿童，父母都不在家，打电话也不方便。有的孩子还是离婚家庭，甚至有的孩子父母都不在，是孤儿，老师出题让孩子回家和父母沟通，是不是也有问题？

2014年北京高考出小作文，题目说：同学们，今天早晨是你的家长送你来

考场的吗？请对家长送考现象进行评论，写一篇不多于 150 字的小作文。

题目里说“你的家长送你来考场的吗”，如果孩子面临家庭变故，没有父母，而题目总是提及父母，就有点太直接，会让孩子不舒服，进而影响他的成绩水平。如果巧妙一点，题目可以设置为：学校图书馆计划向家长开放，于是制定了新的图书借阅规则，这个规则怎么样？你有什么建议？今天晚上回家后，请你和你的家长或朋友说一说，你会如何说？请你把它写下来。

那么这道题可以了吗？很多老师认为可以了，但是我告诉大家不可以。

这道题目末尾说，把你会说的写下来。一旦我们让孩子写下来，我们看到的是孩子的书面语言表达能力，而不是口头语言表达能力。而且我们还要注意，口头语言表达能力不等于口语交际能力，所以这道题根本没有涉及口语交际能力。可是，我们国家的中考题、高考题在考口语交际的时候好像都是这么考的。这些题目真的评价到了学生的口语交际能力吗？没有。学生得分了，我们觉得中国孩子口语交际能力很好，这是不是自欺欺人？有研究表明，中国人口语表达能力不如人意。我们上台讲话一般都要念稿子才能讲清楚，否则就说不清。我们的口语交际能力需要提高，这跟教学有关系，跟教学评价更有关系。因为中高考是那样考的，所以就误导中小学教师那样去教，然而他们教的不是真正的口语交际。

我们必须改进口语交际的评价。其实改进评价并不难。我出这样一道题：同学们，上完“少年爱因斯坦”一课，你怎么看待爱因斯坦对于当时学校教育的不满？如果有同学对我校某些教育方式不满意，你觉得怎么做合适？请同学们分成四人小组，就上述问题进行讨论。

在孩子们讨论的时候，老师会经常在旁边走动观察孩子们讨论的情况。当然，老师们不是在看热闹，而是有目标的。比如在文明得体的交流方面，我发现某个孩子不够文明得体。并不是说他说话粗俗，而是这个孩子表达与众不同的意见时，喜欢冲着脖子，显得不够文明得体，于是我们可以给他一些反馈和建议。这样一来，评价就是有用的。

在改进评价之后，要收集数据和既定目标做对比。如果目标达成了，就可以向前推动教学，如果发现目标没有达成，就需要思考，哪个目标没有达成？为什

么没有达成？怎么样可以更好地达成？然后我们再去补救教学、改进教学，再重复实施评价，直到我们可接受孩子达标的程度。我们指的是可接受的程度，因为让班里的所有孩子都 100% 达成教学目标，再好的学校也做不到。所以达到可以接受的程度了，就可以继续向前推动进度，这就是泰勒模式。

泰勒模式给我们的启示是要坚持科学有效，只有评价能够准确地对学生的习达成目标的程度、对老师达成目标的程度做出准确有效的判断，这样的评价才能发挥其导向作用、诊断作用和发展作用。如果结果评价没有做到科学有效，那么再怎么去加强也没有用。

接下来，我想通过两个例子让大家快速体会为什么要坚持科学有效，以及如何分析结果评价是不是做到了科学有效。

我们先来看一道幼儿园的题目。题目说上面有四只兔子，下面有四辆小车，已知中间那只叫小丽，那另外三只兔子是谁呢？已知小丽在小白的右边，小花在小白的左边，小小在小丽的右边。请你用线段将每只兔子和它的小车连起来。

➤ 请用线将兔子和自己的小车连起来



对于这道题，你明白老师的意思吗？这位老师想了解孩子对于方位概念的掌握程度。那么这道题做到科学有效了吗？它能不能帮助老师判断孩子左右概念的掌握程度？有人觉得不行，孩子是正面区分左右，还是需要转过身来区分左右，这并没有弄清楚。况且是分孩子的左右还是分兔子的左右，这也是很多人的疑问。

假如没有参照点的分歧，孩子都知道以自己为基准。在这样的情况下，这个

题能不能帮我们准确判断孩子会不会分辨左右了呢？有人说应该可以了，我告诉大家不可以。当年我儿子上幼儿园的时候，我把这题给他做，并且他做对了。但由于我是学心理学的，我知道幼儿园的孩子对于左右概念其实没有完全坚定，他平时是会犯迷糊的，可是他做对了。于是我就问，为什么最左边那只叫小花？我儿子说是因为那只兔子脑袋上有小花。为什么最右边的叫小小？他说因为个子最小。我儿子做对了题，那么他会不会分左右？我还是无法判定。那么这道题就不算科学有效，不能准确判断目标达成程度。

我们需要分析评价的效度，也就是评价的科学性、有效性。我们不仅要会分析，也要会对评价进行改进。

举个例子，我们来看一道中考题：一般菊花在秋天开花，影响菊花开花的生态因素是什么？科学兴趣小组对菊花开花与日照长短的关系进行了探究。第一，做出假设。第二，进行实验。取十份大小、长势都差不多且未开花的菊花放在其他环境条件都适宜且相同的地方，其中五盆放在日照长的地方，另外五盆日照短，经过一段长时间后观察开花情况。该实验的主要变量是什么？如果他们假设成立，观察到的现象应该是什么？

➤ 中考生物试题：探究能力评价

一般菊花在秋天开花。影响菊花开花的生态因素是什么？科学兴趣小组对“菊花是否开花与日照长短的关系”进行了探究。

(1) 作出假设：_____；

(2) 进行实验：取十盆大小、长势都差不多且未开花的菊花，放在其他环境条件都适宜且相同的地方，其中五盆放在日照时间短的地方，另五盆放在日照时间长的地方，经过一段时间后，观察开花情况。该实验的主要变量是_____；

(3) 如果他们的假设成立，观察到的现象应该是_____。

这道题考查学生的探究能力，探究能力是生物学科的核心素养，所以题目的立意很好。那么这道题有没有科学有效地评价到学生的探究能力呢？

比如第一小题做出假设，而这里需要孩子做出假设吗？不需要，因为题干说

要探讨菊花是否开花与日照长短的关系，把这个句子转化一下就出假设了。菊花开花与日照长短有关系，或者菊花开花与日照长短无关，或者日照越长，开花越早或越晚。所以这里实际上看不到学生做出假设的能力。

第二小题似乎想要考察学生的实验设计能力，可是实验设计都给定了，还能看到实验设计能力吗？题干说，该实验的主要变量是日照时间，如果我填开花时间可不可以？命题老师认为不行。那难不成在一个实验里边，自变量是主要变量，因变量就是次要变量吗？所以这句话是存在歧义的。

第三小题问如果假设成立，结果是怎样的？但这还是假设，所以这个题没有意义和价值。从整体上来看，如果这道题作为一道中考语文阅读理解题，我认为是一道好题。但是如果作为一道考查探究能力的生物中考题，我就会有疑问，因为我觉得没有做到科学有效。

如果让我来改编这个题目，我会这样来改：一般菊花在秋天开花，通过观察发现在校园里摆放的盆栽菊花的开花有早有晚，请问影响盆栽菊花开花早晚的生态因素是什么？请提出两个假设，如果想到的不止两个，可以多写。

那么假设可以是什么？学生可能会说与气温有关，与空气的湿度有关，与土壤的湿度或肥力有关，还可以与浇水施肥、风力大小、昼夜温差以及菊花品种有关等等。那这么看来这道题是不是太简单了？我想说，在刚才那几个答案里面，有几个我不会给分。比如与菊花品种有关，这是没有分的，因为题目要探讨的是影响菊花开花早晚的生态因素，而菊花品种并不是生态因素。

如果学生写与浇水施肥的情况有关，我也不会给分。因为题干说一般菊花在秋天开花，意思就是只有很少部分的菊花会在夏天或冬天开花。为什么一般的菊花会在秋天开呢？因为秋天风大，所以菊花开花早晚可能与风有关系，或者因为秋天的昼夜温差大，因为秋天一到天就凉了。如果根据天气变化因素来提出研究假设，那么我就会给分。可是写浇水施肥就不能给分了，难道浇水施肥后，那些菊花就会变成冬天或者春天开吗？假设要大胆，但是也有质量，也有好坏。这样一类的题目才具有挑战性，能考查学生基于生物学概念、生物学思想去提出假设的能力，这才是真正的探究能力。

如果要考查实验设计能力，我就会问，哪些生态因素会影响菊花开花早晚？有人假设日照时间是影响菊花开花早晚的一个因素，于是他设计了如下实验——取十份大小、长势都差不多且未开花的盆栽菊花放在其他环境条件都相同的地方，其中五盆放在日照时间短的地方，另外五盆放在日照时间长的地方，经过一段时间后观察开花情况，然后出了3道小题。

这个实验设计好不好？能否能有效验证假设是否成立？考查实验设计能力一般有两种办法，一是基于问题假设提出实验设计；二是有假设也有实验设计。我会选择第二种，让学生去评价实验设计。

很多老师说这是生物中考题里的实验，是当地最优秀的老师命制的中考题，怎么会有问题呢？这道题用了控制无关变量的方法，这是很好的。但是如果让我来评价这个实验设计，我认为它不合格。它存在许多问题，比如样本量才五盆，这是不够的。

举个例子，我想知道是天津市和平区的老师普通话说得好，还是重庆市北碚区的老师普通话说得好。我在和平区和北碚区各选五名教师，考查他们的普通话，就此下结论，您觉得是不是不靠谱？样本量、组数是不是不够？实验就两组，也没说清楚间隔，那么是不是可能都开或者都不开？这看不出来日照时间的影响。最重要的是，这个实验似乎控制了无关变量，但是没有控制菊花的品种。连品种都没有控制，即便真的有组菊花先开了，有组菊花晚开了，那是受光照时间影响还是菊花品种影响呢？这是不是就说不清楚了？这样来做实验是不行的。

如果所有生物学老师命制的实验都有这么多问题，孩子的探究能力是不是很难提升？所以我们国家要建设高质量的教育，要使孩子真正具备迎接21世纪挑战的核心素养，我们的评价就必须改变，包括教学评价、结果评价都要改变。

从这个意义上来说，老师要加强教学结果评价的研究，既要是教学能手，又要是评价专家。我们的教研活动研究教学好像研究得很透彻了，但是只评价不研究是不行的，因为只有老师会评价才能真正会教。2016年我出版了一本书叫《促进教学的测验与评价》，里面有很多例子，老师们可以看一看。我们需要思考如何一边测验一边评价，这和设计作业是一样的道理。

接下来我们看看对**教学的过程性评价**。

我们在实践中要做好过程性的教学评价，而过程性的教学评价主要是通过听评课来进行的。那我们要怎么去听评课才能做好过程性评价呢？我想先带大家听一小段公开课。

一位老师几年前上了一节公开课，这节课讲的是司马光，那时候还没有用部编、统编教材。老师讲了一个关于司马光的故事，故事里说司马光小时候有一次在后花园玩，可是一个小伙伴特别淘气，不小心掉水缸了，而且水缸里的水特别深，要是再不把他救出来，就有生命危险。在这个危急时刻，司马光应该怎么办呢？有孩子举手说，拿一个吸管一样的长棍，让那个小朋友先呼吸，再想其他办法救人。老师说：“你是个善良且充满智慧的孩子，可是要找一个像你说的木棍可能来不及了，同学们还有别的办法吗？”

第二个同学举手说，可以找一块大木板，让小朋友爬上去浮起来。老师说：“你真爱动脑筋，知道的知识也不少，是个聪明的孩子，可是到哪去找这么一块合适的木板呢？还有，水缸是不是足够大？这都存在问题，谁还有别的办法？”

第三个同学举手说，找一个软的皮管子，把一头放水缸里，另一头用嘴吸水，这样做小朋友不也得救了吗？老师说：“你真是个细心的孩子，这么善于观察，现在又学会应用了，真了不起，可是要在当时找到我们所需要的皮管子可能也来不及了。”在当时的情况下，司马光到底该怎么办才好呢？请同学们翻开书到书中去找答案。这就是这位老师公开课课堂导入的片段。

在这节课里，我们很容易能发现这位老师注重师生互动，希望将教学的过程变成师生互动、对话交往的过程，希望孩子主动参与到课堂学习里来。可是您会不会觉得他的互动有点假？因为学生课前有预习，所以他们应该知道砸缸。即便课前没有预习，难道班里的孩子真的不知道砸缸吗？不用说城市里的孩子，就是偏远山区里的孩子，我也觉得至少应该有一两个孩子能说出砸缸来。可是您看，班里就没有一个人说砸缸，他们都在想其他办法，说明他们都在配合老师，想老师之所想，师生通过配合达成了—个虚假的繁荣而已。那么在这样的课上，学生只是配角，他们的主体性就会丧失，而让学生成为学习的主人、以学定教、因材

施教的课堂教学理念也只能落空了。

这样的课是不是很多？有校长和老师跟我说，他们早就改了，不会再这样。的确，或许学校里这么夸张的虚假互动已经消失，但是它的影子仍然在。我经常去学校听课，还是会看到这些现象。

那天我去听一个物理老师讲课，讲的是滑动摩擦力。在讲课之前，我还跟孩子们聊天，因为我想知道孩子们课前预习了没有，了解了什么，还有什么样的困惑，由此来判断这个老师的教学设计是不是以学定教、因材施教。

班上每四个同学坐在一块，我就知道老师在课上会有小组探究活动，设计应该还挺新颖的。我跟这四个孩子聊天，我问第一个孩子，今天讲什么内容？孩子说今天讲滑动摩擦力。我又问能不能说得具体一点？他回答，讲滑动摩擦力的大小跟什么影响因素有关。

我问了第二个孩子，我想看看他预习的情况。第二个孩子很能干，他张嘴就来，滑动摩擦力大小肯定跟物体重量有关，物体越重，滑动摩擦力就越大，除此之外还和接触面有关。在他还要继续说的时候，坐他斜对面的孩子在桌底下踹了他一脚，然后他们几个就不说话了。我又问第四个孩子，滑动摩擦力还和什么因素有关？我想让他补充一下。可是第四个孩子“政治觉悟”特别高，他满脸严肃地说不知道，老师还没有讲过。可是他真的不知道吗？他其实也知道了，因为他课前有预习，甚至老师为了上好公开课，在课前已经把这节课给捋一遍了。可是他们现在突然意识到，老师的公开课还没上，专家问你，结果你都知道，这怎么行？所以要假装不知道。

后来，等到再上公开课时，这个老师仍然假装孩子不知道。他说：“同学们，今天我们讲滑动摩擦力，我们要探讨的问题是滑动摩擦力的大小跟什么影响因素有关？它究竟和什么有关呢？你们能不能先提出一些假设，想想可能会和什么有关？”于是孩子们开始举手，有人说可能跟物体的重量有关，有人说可能跟接触面有关等等。孩子们配合得特别好，他们不仅提出他们知道能成立的假设，还假设出一些不靠谱、不可能的答案。既然是假设，绝对不能全是正确的，也得要装装样子。

那么究竟这个滑动摩擦力和什么影响因素有关呢？老师在讲桌上放了一个东西，然后用弹簧拉力计以匀速直线运动来拉它，这个时候拉力计的示数就是滑动摩擦力的大小。然后这位老师掏出来一块浴巾铺在讲桌上，又把刚才那东西放在浴巾上。他说现在我要再去拉它，你们说滑动摩擦力会怎么变化？孩子们继续配合，说摩擦力会变大。

于是老师就继续拉动物体，孩子们又很配合地去观察这实验，结果确实是变大了。一位同学说：“这说明滑动摩擦力大小跟接触面有关。”另一位同学说：“他说得不够完整，应该是滑动摩擦力的大小跟接触面的粗糙程度有关。”

您看，孩子们在老师的带领下，通过实验探究发现了一个真理。可是其实孩子们老早就知道这个答案了，不用说这些孩子们，就算是一天学也没上过的老农民都知道这个答案，可是老师假装孩子不知道，然后还设计探究，接着恍然大悟。

您觉得这节课真的发生了吗？学生有没有深度学习？学生有没有实际获得感？这其实都存疑，因为学生在配合老师。为什么这种现象很多？原因是多样的。我们总是提倡要上优质课，要磨课。但是我们可以在不同的班、在没上过的班来磨课，但是上过课了还假装磨课，那不就是表演了吗？这对学生是有伤害的，也是不道德的。

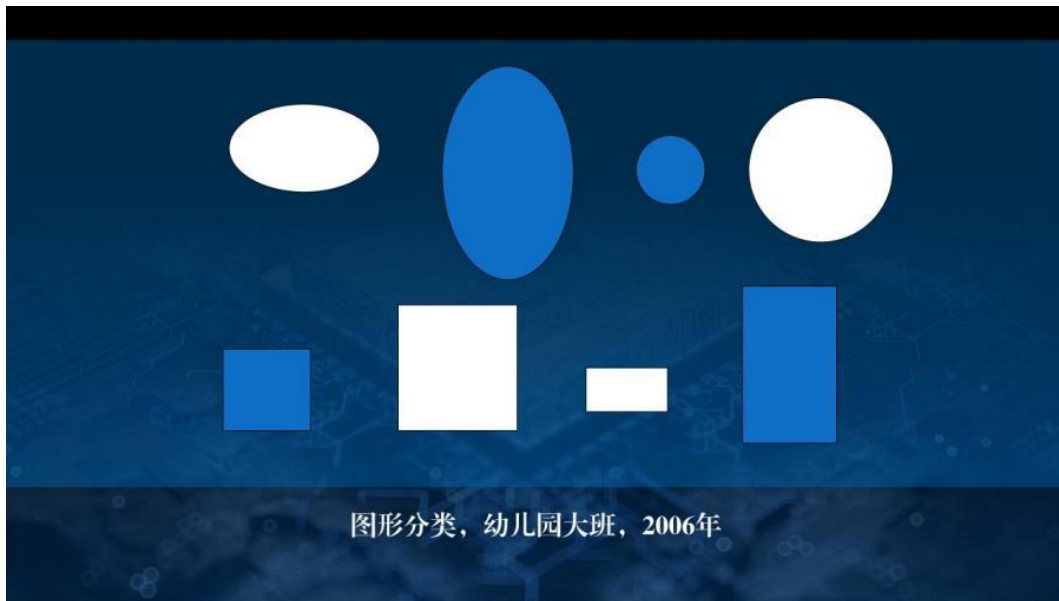
如果这种虚假的课堂没有改变，孩子的主体性会受到莫大的伤害。

比如，我们可以看下图。这是我儿子上幼儿园的时候老师布置的家庭作业。老师让孩子们给八个图形分类。我大儿子看了看说，可以把长方形和正方形放一边，把圆形和椭圆形放一边。我问还有别的方法吗？我鼓励孩子们做发散性思考。

这时候二儿子看了看，说有的图形有颜色，有的没颜色，可以把有颜色的放一边，没颜色的放一边。我继续鼓励他们，说：“挺好的，还有没有别的方法呢？有第二种说不定就有第三种。”

我的二儿子又继续说，圆形是“双胞胎”，它们是哥俩，一大一小，椭圆形、长方形、正方形也是一样，可以把“双胞胎哥哥”放一边，把“双胞胎弟弟”放一边。我说太好了，我发自内心地赏识孩子，因为这是孩子自己发现，是用自己的语言描述出来的。可是我话音才落，我的大儿子就问我，老师认为哪个对？我

听了以后，心凉了半截。



透过这个发问，我们会发现孩子虽然还小，但是他的主体性以及和主体性相关的独立思考意识、创造性想象力、批判性思考能力等等，这些新课程倡导的核心素养已经被老师祸害得差不多了。

我们常说中国人缺乏创造性是由中国传统文化导致的。可是现在想想，我们身边的孩子缺乏独立性、创造性、想象力、批判性、自觉性，缺了这么多核心素养，是不是和每一个老师的课息息相关？在我们的课堂上，提出问题之后，孩子只会想“老师想让我说什么”“老师认为哪个对”“老师是怎么想的”，可他没有问真理是什么？我对这件事有何看法？我想如何解决问题？这样一来，何谈主体性，何谈创造性？所以我们的课堂必须改变，**我们必须把课堂还给学生，必须让学生成为主体**，让学生真的去跟老师对话、交往、互动，这样才能有真正的学习发生，所以这点很重要。

除了这一点，刚才“司马光”一课中，还有一个问题值得我们讨论。我们注意到，刚才这位老师在孩子回答完问题之后都及时地反馈、鼓励和表扬。不论是不是上假课了，我们都需要看到他的鼓励表扬。假设，刚才上“司马光”这一课的学生都不是中国人，而是美籍华人，美国生，美国长，只是到中国大陆参加个夏令营而已。老师提完问题，孩子提想法，然后老师就及时鼓励、反馈、表扬，给予评价指导。孩子们真不知道砸缸，老师也没有跟学生做虚假互动。在这样的

情况下，您觉得这个老师的评价、反馈、表扬好不好？有老师会说挺好的，但我想说不好。因为我们发现这位老师虽然表扬了孩子，可是他最后都是否定态度。

心理学家罗杰斯说，每个人最需要别人给他的是无条件的赏识，这样他才能够自我接纳，才能有真正的自信心和活力。可是现在老师和家长，往往给予的是条件性的肯定，比如“你做得好，但是有两个小小的问题”。您觉得这样的话是在鼓励表扬孩子，还是在否定孩子呢？本质是否定的。罗杰斯认为这种条件性肯定给多了，会使孩子产生紧张感。

比如在一堂英语课上，老师在讲完“How’s the weather?”后会让两个学生进行模拟对话。第一组同学站起来一问一答，老师说：“你俩说得挺好，可是这个词有点问题，再跟着老师说一遍就更好了。”虽然老师面带微笑，眼里放着爱的光，可孩子能体验到这份爱吗？老师上课总是拐弯抹角，找个毛病才让孩子坐下。孩子就会觉得早知这样就不举手了。

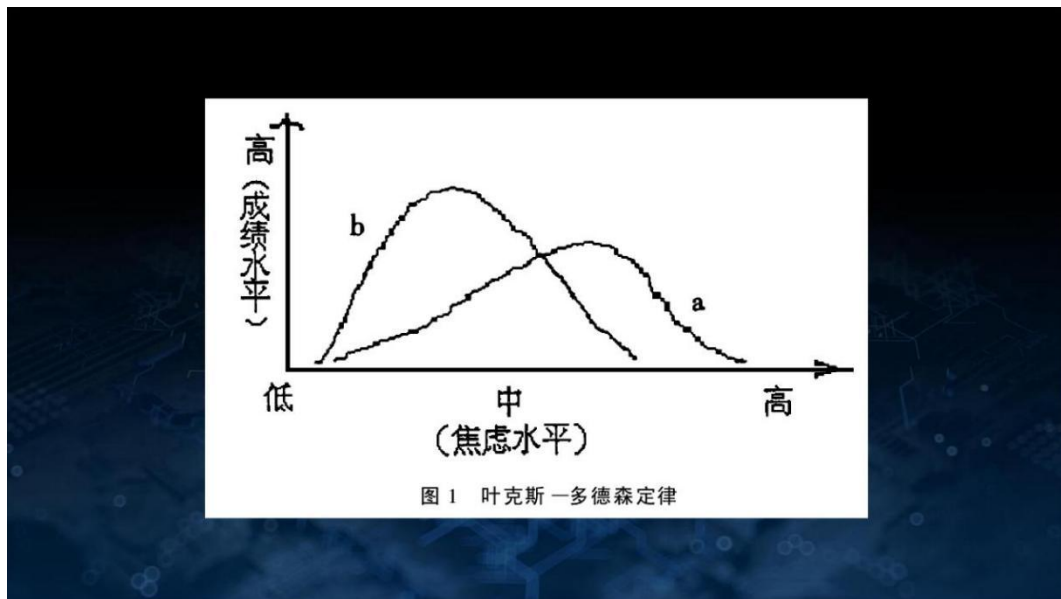
之后，第二组同学站起来说完了，老师说：“你俩说得还行，接下来我们看谁能说得比他们更好。”这俩孩子就觉得自己也说得不好，所以他俩也灰溜溜地坐下来。等第三组同学说完，老师说：“你俩说得真好，口音真好听。”那这样一说，前面发过言的孩子都明白意思了。

虽然，这个老师可能课上上面带微笑的，但是孩子们都没有体验到无条件的积极肯定。如果是公开课，或许有人配合，举手回答问题。倘若不是公开课，那谁还会搭理他？为什么孩子学英语，学到最后变成哑巴英语？就是因为他身边的人总是给予条件性肯定，没有真正赏识过他。

当然，有老师会心生疑问说：“如果孩子语音不对，语法不对，我们就不说吗？如果老师教数学、教物理，孩子错了，到底该不该说？”

叶克斯提出了一个定律，一个人的焦虑和一个人的成绩不是线性关系，不是说学生越焦虑，成绩就会降低，这两者呈现出的是倒U型关系。A同学那条线在一定范围内，越焦虑，成绩会越提高，但是当焦虑达到一定程度，他的成绩会下降，B也是如此。所以我们要注意，适度的焦虑、适度的压力、适度的紧张感对学生是有积极促进作用的，所以孩子在成长中是需要否定，需要指出问题的。

但是如果孩子已经到了曲线的后半段，已经特别紧张、特别焦虑了，如果再给他增加压力、增加紧张感，他更没自信，更不能好好学习了。这都是因情况而异的。



我们再来看这张图，A和B两位同学的曲线不一样？这叫个体差异。孩子跟孩子不一样，同样的问题，同样的焦虑水平，给A可以，但是给B就不可以。这个问题都做错了，您批评A可以，但是不能批评B。我们需要看情况，而且说的时候要讲究方法，不一定是“你真好”，真诚地直接说有时比绕弯子还好。

大家有空可以做一个家庭实验，任务是指导某人打电话。假设你选择以你的爱人为实验对象，你不能一开始就说这不好，那不好，得先说优点，再说缺点。等他把电话打完了，你可以说，你的语言能力真强，有三个优点，但是还有三个缺点。优缺点都要分点说清楚。打两到三个电话，实验就可以结束了。那么实验结果会怎么样？或许两口子会打起来，也可能你爱人很有风度，很温和，没说什么。但是你会发现一个微妙的变化，以后您爱人电话一响起来，他抓起手机就往房间里跑，关上门再接电话。

条件性肯定的威力特别大，可是我们的孩子就活在条件性肯定中。而且，父母是老师的孩子最悲惨。因为老师会把自己的孩子跟优秀的孩子比，天天都这样，孩子受得了吗？

通过司马光这一课，我们发现有很多需要分析改进的问题。如果把过程性教学评价做得扎实、深入了，真的可以发挥导向、诊断、发展作用。当然，要发挥

好，除了时间花得够、做得足，还要注意一些理念性的问题。在我们的实验学校做教学评价，我们强调以学评教，也就是说把教学评价的焦点放在学生的“学”上，而不是教师的“教”。我们不是看教师的教态是否自然，普通话是否标准流利，教学设计是否循序渐进，教案写得是否细致工整。我们的重点是学生的学习真正发生了吗？学习的深度够不够？他们掌握了应该掌握的知识了吗？他的核心素养发生变化了吗？情感态度上有什么样的体验？是积极愉悦的，还是需要去改变提升的？这些才是判断老师教学好不好的最重要的标准。学生情绪开不开心？受到鼓励没有？以后是不是能够自主学习？知识有没有掌握？这是我们判断好坏的标准。

另外，教学评价不是“就课论课”，而是“以课为例”“以课为镜”。我们每个人都拿这堂公开课来照镜子，这很重要。如果我们总是就课论课地说，小张老师的课有三个优点，有三个缺点，那谁也不愿意上公开课。一上完公开课就要接受别人的批斗，因为无论上得多好，都可以找出问题来。虽然说我也愿意成长，但是一般人很不愿意接受这样的挑战。

可是，如果我们仅仅以张老师的“司马光”这节课作为一个课例、一面镜子来反思教学，研究教学如何改进，是不是就不一样了？现在我们可以说，听了张老师的课后，有两点收获，每一点收获实际上对应的是张老师的优点。但是倘若我说张老师有什么优点，好像是自上而下的评价。如果换成课堂收获的说法，一个平等互动的研讨氛围就形成了。大家透过课堂去看教学中的问题，去思考如何共同改进教学，这才是更有意义的教学评价。

无论是结果性教学评价还是过程性教学评价，我们都需要重视。对于实践中的问题与挑战，我们需要不断反思，不断改进，一起努力。



校长内参

一个好校长就是一所好学校



明德云学堂

大咖零距离 尽在明德云