



北京大学医学部  
教学发展中心

# 课程考核

蔡景一

北京大学医学部教育处/教学发展中心

2014年12月13日



# 内 容

- 为什么要做课程考核
- 如何设计课程考核
- 如何做试卷分析
- 如何评价课程考核



# 内 容

- 为什么要做课程考核
- 如何设计课程考核
- 如何做试卷分析
- 如何评价课程考核



# 为什么要做课程考核

➤ 课程考核是课程计划的一个重要环节

- 了解学生差距
- 鼓励督促学生进步
- 确保学生达到标准
- 提供达到毕业或获得学位的成绩
- 提示重要的课程内容



# 内 容

- 为什么要做课程考核
- 如何设计课程考核
- 如何做试卷分析
- 如何评价课程考核



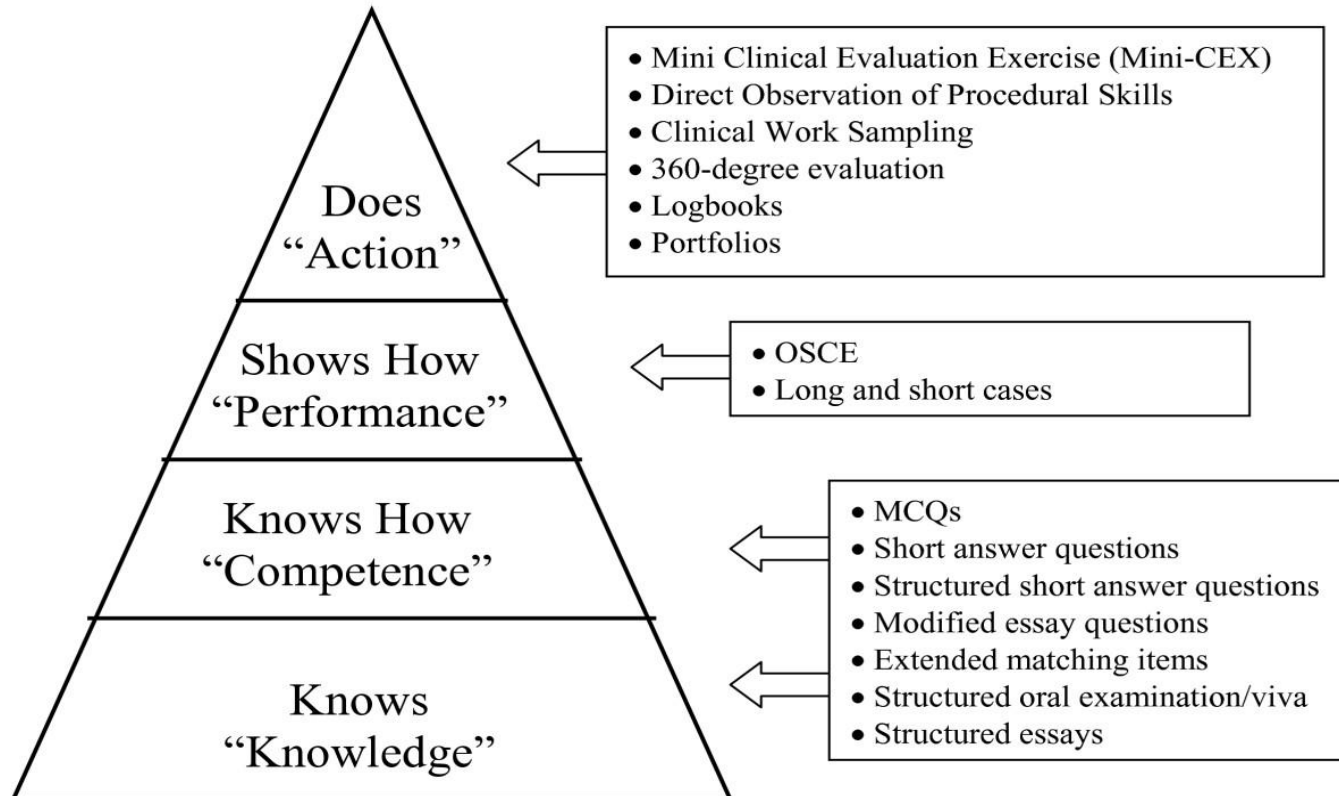
# 如何设计课程考核

- 以学习目标或以结果为基础
- 符合课程教学大纲要求
- 内容
  - 知识领域：记忆和理解能力
  - 技能领域：精神运动任务，职业所需
  - 专业态度：职业所需的个人品质
- 评价工具
- 评判标准确定
- 考核结果分析总结



# 不同内容：不同考核方法

**Knowledge < technical skills < non technical skills < attitudes**





# 临床常用的考核方法

## MCQ

Multiple-choice  
questions  
多选题考试

## Mini-CEX

Mini-Clinical  
Evaluation Exercise  
迷你临床演练评估

## OSCE

Objective Structured  
Clinical Examination  
客观结构化临床考试

## ITER

in-training evaluation  
report  
训练评价报告

## MSF

Multi-source  
feedback  
多源反馈

## Portfolio

文档与日志记录

## DOPS

Directly Observed  
Procedural Skills  
操作技能直接评估

## Short Case

短病例考试

## Long Case

长病例考试





# 评判标准设置

- 标准设置目的是确定一个科学的划界分数，区分考生中合格与不合格/平均水平的过程。
- 相对标准:适用选择一定数量或者比例的考生的评价，如录取或录用测试。
- 绝对标准:适用于确定考生是否掌握必需的知识或技能的评价，如课程考试、毕业考试等。
- 标准的设定依赖于主观判断。



# 基于试题的评判方法

- 最常用的两种方法：Angoff法和Ebel法
- Angoff法：直接判断边界组考生在每道试题的正答率
- Ebel法：先将试题按照一定标准分类，再估计边界组考生在每类试题的正答率
- 边界组选择



# Angoff法

问题	标准设定者					专家百分数
	A	B	C	D	E	平均值
1	.90	.85	.80	.75	.85	.83
2	.60	.55	.40	.35	.50	.48
3	.70	.60	.65	.50	.55	.60
4	.85	.75	.80	.65	.70	.75
5	.95	.90	.85	.75	.80	.85
6	.50	.50	.45	.40	.50	.47
7	.65	.55	.45	.45	.60	.54
8	.85	.70	.80	.65	.75	.75
标准						平均正答率



# 内 容

- 为什么要做课程考核
- 如何设计课程考核
- 如何做试卷分析
- 如何评价课程考核



# 试卷分析

- 题型结构及分值分布情况
  - 主观题——问答题(简答题、论述题)、分析题等
  - 客观题——选择题、扩展配伍题等
- 试卷的信度、效度分析
- 难度、区分度分析



# 试卷信度

- 信度：代表测试的稳定程度。
- 试卷信度分析
  - 分半信度：测量考试的内部一致性。
- 影响信度的因素
  - 评分的客观性：对评分标准的把握
  - 环境影响：如不同考试时间
  - 偏倚误差：考官有偏见、关注某一特征



# 试卷的效度

- 效度：代表测试的准确性，所测试的内容是否是应该测试？
- 试卷效度分析
  - 内容的覆盖面：涵盖需要掌握的知识点
  - 认知能力的分级：记忆、理解、分析、应用等
  - 题型选择：多选题、简答题、论述题等
- 影响效度的因素
  - 试题表述含糊或具有误导性
  - 试题数量太少
  - 答题时间不足
  - 试题的难度
  - .....





# 生理心理学MCQ的命题方案

内容	记忆	理解	分析	应用	合计
神经元/解剖	6	4	2	1	13
运动机制	4	3	2	1	10
感觉系统	3	2	2	1	8
预警机制	3	2	2	1	8
动机/情感	2	2	1	1	6
记忆/学习	2	2	1		5
合计	20 (40%)	15 (30%)	10 (20%)	5 (10%)	50 (100%)





# 难度分析

## 预设标准与实考难度的比较

题目	标准设定者					平均值	
	A	B	C	D	E	预设正答率	实考正答率
1	.90	.85	.80	.75	.85	.83	0.90
2	.60	.55	.40	.35	.50	.48	0.50
3	.70	.60	.65	.50	.55	.60	0.70
4	.85	.75	.80	.65	.70	.75	0.70
5	.95	.90	.85	.75	.80	.85	0.80
6	.50	.50	.45	.40	.50	.47	0.60
7	.65	.55	.45	.45	.60	.54	0.45
8	.85	.70	.80	.65	.75	.75	0.50
试卷						0.66	0.64



# 内 容

- 为什么要做课程考核
- 如何设计课程考核
- 如何做试卷分析
- 如何评价课程考核



# 如何评价课程考核

## ➤ 总体情况

- 考核内容是否符合教学目标和教学大纲要求
- 评价方法和评判标准是否合适

## ➤ 学生成绩分布情况

## ➤ 试卷信度、效度和难度分析

## ➤ 存在不足及改进意见

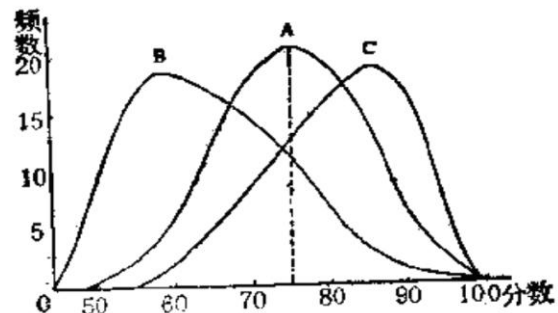


图 10-1 分数的频数曲线图

A 正态曲线    B 正偏态曲线    C 负偏态曲线



# 评价结果的使用

- 评价是指挥棒
  - 想要学生学什么？
  - 要达到什么水平？
- 评价结果应及时反馈给学生
- 评价结果应用于改进教学



北京大学医学部  
教学发展中心

# 我们的团队







北京大学医学部  
教学发展中心

欢迎各位老师关注北京大学医学部教学发展中心

Center for Excellence in Teaching and Learning. PUHSC

微信公众平台

网站: <http://tl.bjmu.edu.cn>

邮箱: [tl@bjmu.edu.cn](mailto:tl@bjmu.edu.cn)



感谢您对我们工作的关注与支持!