

# 高校课程考试基础知识



李文国



CONTENTS

目录

- 1 考试概述
- 2 命题与制卷
- 3 测后分析
- 4 考试改进

Overview of the examination

# 考试概述

# 01



# 一、考试概述

## (一) 考试



# 一、考试概述

## (二) 类型

01

安置性考试（前测）

分流

分班

分组





# 一、考试概述

## (二) 类型

02

诊断性考试

— 发现问题

— 改进教学

不作为学生成绩依据

供教师教学参考





# 一、考试概述

## (二) 类型

03

成就性考试

阶段性

课程

综合

04

合格性考试 ( 毕业考试 )



# 一、考试概述

## (三) 方式

### 1. 闭卷（笔试）

- 考查学生对基础知识基本理论的记忆、理解、分析和简单应用
- 题型多、覆盖面大、客观题占比大
- 题型：填空、选择、判断、名词解释、简答、计算、论述、案例分析等

### 2. 开卷（笔试）

- 考查学生对知识的理解和综合应用、学习能力和查资能力

### 3. 口试（答辩）

- 面试；有深度、欠客观





# 一、考试概述

## (三) 方式

### 4. 成果考试 (设计、论文、报告、制品)

- 书面表达 (查阅、计算、绘图、制作能力)
- 或实物作品 (说明)
- 为保证其真实性应与口试结合

### 5. 操作考试

- 考查学生技能, 对主试要求高

### 6. 观察考核

- 用于学生职业素质考核



# 一、考试概述

## (三) 方式

### 7. 标准化考试

- 命题标准化、答案标准化、施测标准化、评分标准化、计分标准化、分数解释标准化。
- 量尺标准、程序科学、误差可控、信效度高；淡化人为因素、客观、自成系统

### 8. 无纸化考试（发展趋势之一）





# 一、考试概述

## (四) 考试方式设计

### 1. 基础学科

闭卷考试  
≥60%总成绩

口试、操作  
40%

### 2. 专业课

开卷笔试  
60%

闭卷考试  
40%

### 3. 语言类

笔试、口试相结合

Propositions and making volumes

# 命题与制卷

# 02



## 二、命题与制卷

是教师的基本技能之一

### (一) 原则



### (二) 计划 (双向细目表)

认识层次 知识内容	按布鲁姆教育目标分类或按题型 认知领域：记忆、理解、应用、分析、综合、评价	合计
章节 知识 内容		
合计		100

## 二、命题与制卷



### (三) 质量要素

#### 1. 试题编写

- 科学性
- 依据课程教学大纲或考试大纲
- 在双向细目表中体现

#### 2. 难度控制

#### 3. 误差控制

- 主观题的误差控制是关键，要防止测验冒犯或不公平惩罚

## 二、命题与制卷

### (三) 质量要素

#### 4. 题型

国内

主观题

客观题

国外

建构性反应试题

选择性反应试题





## 二、命题与制卷

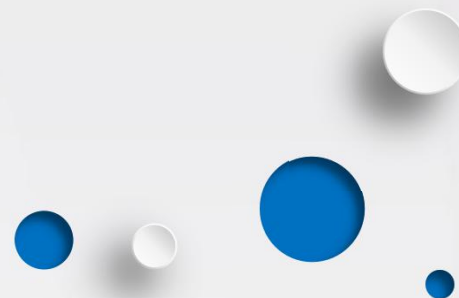
### (三) 质量标准

#### 5. 题量

• 学生用时：教师用时 = ( 2~2.5 ) : 1

#### 6. 优化试卷

- 信度：试题覆盖面要大，题量不能太少，选择项多的选择题
- 效度：先易后难，不同试题内容不能互相牵连
- 难度、区分度：同质选择项越多试题越难

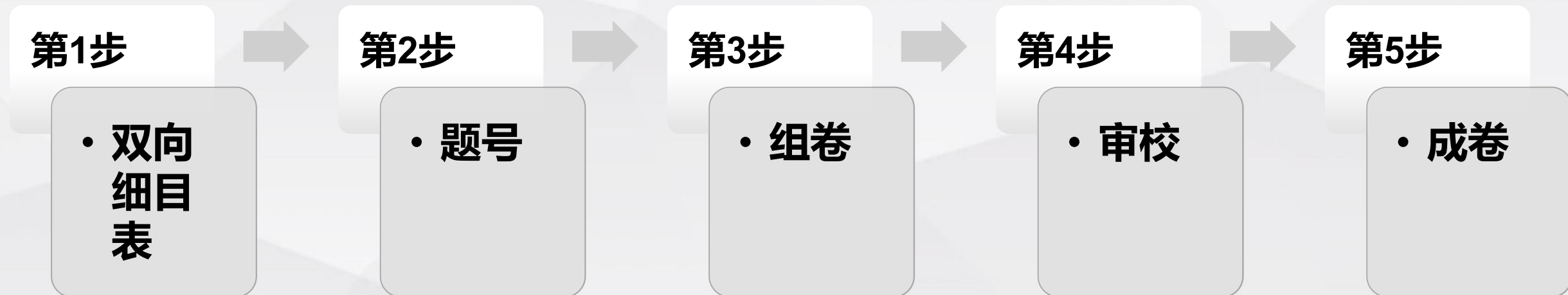




## 二、命题与制卷

### (三) 质量要素

#### 7. 编卷



Examination paper analysis

测后分析

03

# 三、测后分析



## (一) 试卷分析：基本环节 常规工作

### 1. 定性分析

- 整体性分析：成绩分布特征，因素分析
- 科学性分析：试题内容、语言表述、插图规范、答题要求
- 一致性分析：重点与覆盖面、知识与能力、理论与实践

### 2. 定量分析

- 总分均值，标准差

## 3. 要素分析

- 试卷的知识、能力要素结构是否合理
- 知识、能力分数、权重分配是否合理
- 题型结构是否合理
- 内容是否具有代表性（样本）
- 试卷长度和测试时限是否控制良好
- 难度
- 区分度
- 考试结果是否具有-定的稳定性（信度）、可靠（效度）

## 4. 改进建议



# 三、测后分析

## (二) 考试质量评价

### 1. 信度

- 多次测量的一次性
- 考试实得分数与学生真实水平是否吻合，可信
- 信度系数 $0\sim 1$ ， $\geq 0.9$
- 稳定性信度、复本信度、内在一致性信度
- 效度高则信度高，反之则不然
- 考试异常重要时才计算信度



# 三、测后分析

## (二) 考试质量评价

### 2. 效度 (最重要的指标)

- 在一次测验中测量到它所要测量的东西的程度。
- 内容相关效度，效度系数 $0\sim 1$ ， $0.4\sim 0.7$
- 为保证考试分数推断的精确性，多种效度共同使用
- 内容相关效度 (代表程度)，效标关联效度，结构相关效度



## 三、测后分析

### 3. 难度

- 试题的答对率,  $P = \frac{R}{N}$ ,  $P = \frac{\sum Xi}{NA}$
- 理想值  $P = 0.5$

(0.3~0.7)

占70%

(0.2~0.3) (0.7~0.8)

难易题各占15%

(0~0.1) (0.9~1)

特难特易占10%



## 三、测后分析

### 4. 区分度

- 对一道试题，水平高能力强的学生得分高，反之得分低
- $D = P_H - P_L$ ； $P_H$ ：高分组答对率 $P_L$ ：低分组答对率，各取27%

0.3~  
0.39

良好

$\geq 0.4$

非常好

0.2~  
0.29

勉强合  
格

$\leq 0.19$

差





## 三、测后分析

### (三) 成绩分析

#### 标准参照评价

#### 绝对测量

- 对内容（样本）掌握的程度（及格线 $\geq 60\%$ ）

#### 常模参照评价

#### 相对测量

- 个体在群体中所处位置，分数需转换



# 三、测后分析

## 1. 考试分数分布

- 频数分布（图或表）
- 频数：同一分数段中分数个数，如80~89分数段中3个
- 分布：全体考生分数落在各分数段的情况

### 集中量数

- 众数（频次最多的分数）
- 中位数（居中分数）
- **平均数**

### 差异量数

- 全距（最高最低分之差）
- 四分距（把排序后的分数按个数四等分，第三个分点与第一个分点分数差的一半）
- **标准差**



## 三、测后分析

•2.

• 平均分  $\bar{X} = \frac{\sum X_i}{N}$

• 标准差  $\sigma = \sqrt{\frac{1}{N} \sum (X_i - \bar{X})^2}$

• 标准差系数  $w = \frac{\sigma}{\bar{X}} \times 100\%$ , 越小越好



## 三、测后分析

### “标准参照性”考试

- 按课程类型制定相应的考试大纲，向学生公布

#### 选择性反应

- 从提供的答案中选答案
- 设计难 评分易

#### 建构性反应

- 学生独立解决问题 设计易 评分难



# 三、测后分析



## (一) 选择-反应测验题

- 可在更大范围内考查学生对知识的掌握情况



指导语忌语言晦  
涩难懂

忌模棱两可的陈  
述：红花布袋、  
父在母先亡

忌无意中给出线索

忌使用复杂句型



## 三、测后分析

### 1. 判断题（判对 错）

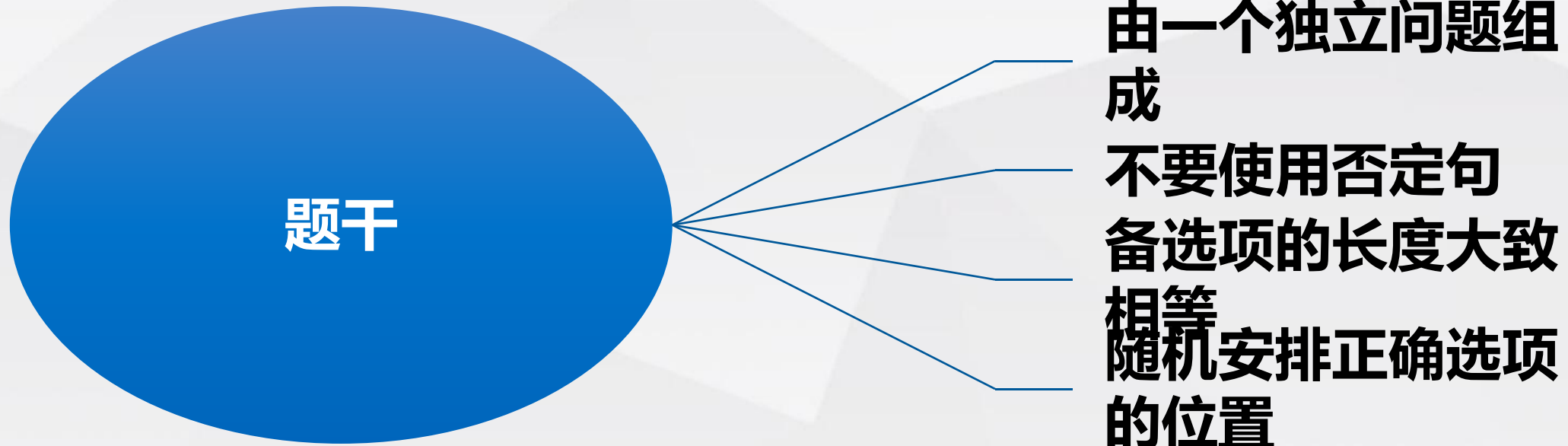
可大题量，可猜测（50%）

- 语言要严密
- 每个陈述只考查一个概念：曾国藩是湖南人
- 少用否定陈述：曾国藩不是湖南人
- 答案“正确”或“错误”题数大致相等，长度大体相当

## 三、测后分析

### 2. 多项选择题： 题干 包含正确答案的备选答案4-5个选项

问句形式 不完整陈述句形式 1个正确答案 其余为干扰项  
如：曾国藩是哪个朝代的人？ 1.元朝；2.清朝；3.明朝；4.南北朝



## 3. 匹配题：

**两列单词或短语、将两列相关信息对应起来**

前提

反应

( 上限为10 , 后者数量应大于或等于前者数量 )

- 各列的同质性 ( 重要属性 )
- 较短的单词或短语放在右边
- 反应应比前提多
- 有逻辑地安排反应的顺序
- 说明匹配要求及每个反应可使用的次数





# 三、测后分析

指导语：左边的A列为战争名，右边的B列为美国总统。请从B列中选出每一次战争结束时的美国总统，将相应的序号填在A列合适的括号中。注意B列中的每个名字最多使用一次。

A 列	B列
( ) 1.第一次世界大战	A 布什
( ) 2.第二次世界大战	B 克林顿
( ) 3.朝鲜战争	C 艾森豪威尔
( ) 4.越南战争	D 约翰逊
( ) 5.海湾战争	E 尼克松
	F 罗斯福
	G 杜鲁门
	H 威尔逊

# 三、测后分析

## (二) 建构性反性测试

**1. 简答题（填空题）** 为一个直接的问题或一个不完整的陈述回答或补充一个单词、一个短语或一个句子。若加长反应即为主观题、否则就为客观题。

- 考查知识掌握情况
- 学生需要自己产生一个正确答案
- 反应越长越难准确评分



## 三、测后分析

### 试题为问句或不完整陈述

反应的简洁性 01

02 留出空白

不要多于两个空格 03

04 空格长度相等



## 2. 主观题 一种特殊形式的写作题（一段或几段文字）

- 考查学生综合、分析、评价能力，层次高
- 对学生的反应进行稳定而可靠的评分难

- 限制性反应题：例如，要求学生从三个方面描述，每种描述不超过100个字
- 扩展性反应题：按照首先、其次、再次、最后的要求描述

### 主观题评分：

- (1) 整体评分法或分项评分法，通常先作整体评分然后对低分者分项评分
- (2) 制订一个评分要点
- (3) 对写作技能的重要性作出决定
- (4) 同一题目评完后再评下一个题目
- (5) 尽量匿名评价

## 三、测后分析



### 表现性评价

- **学生行为—任务 针对重要技能**
- **类推的准确性**

### 档案袋评价

- **个人作品的系统收集：摄影、绘画、记者、模特、建筑师**
- **缺点：耗时**
- **优点：注重学生自我评价**

Examination evaluation

考試改進

04

## 四、考试改进

**“教学决定考试，考试促进教学”，“以考促教、以考促学”。问题：**

- **划考试范围、做模拟题（真题），考试目的（满足公众问责，满足学生学习改善）不清**
- **考试信息反馈机制缺失（在高校尤其明显）**
- **考试评价维度单一：  
重知识记忆、轻分析推理、综合运用**
- **考试命题质量有待提高**

# 四、考试改进



## 考试改进须注意

- 如何避免学生抄袭
- 随意降低教学要求(降低考题难度)
- 迎合学生，使考试缺乏严肃性

## 笔试

- 可以从更高层次考查表达能力，评价过程更为公正，有共同的标准比较学生的水平，可更为有效地检查教师的教学质量和课程质量。

## 终结性考试

- 更关注学生较长时间内的学习进展，督促学生重视长期学习不断温习已经学过的内容，督促学生将分散的课堂教学融会贯通，发展其归纳概括能力。





谢谢！