



如何评估线上学习效果

邱意弘

上海交通大学

教学发展中心

2020.05.013

课堂表现 10%

作业 20%

期末考试 70%



需要调整吗？怎么调？



目录

1. 评估的目的及原则
2. 评估方法1（测试）
3. 评估方法2（表现型评估）
4. 表现型评估的工具—量规 (rubric)





1. 评估的目的和原则



评估的目的

- A. 选拔合适的人才
 - B. 为学生评定等级
 - C. 诊断学生的优势与不足
 - D. 监测学生的发展
 - E. 判断自己的教学效果
 - F. 提高教学质量
- 总结性评估
- 形成性评估
- 自我反思与改进



图片来自网络

1. 学习成果为“记忆知识点”，评估方式可用：

A. 考试

B. 口试

2. 学习成果为“会操作实验仪器”，评估方式可用：

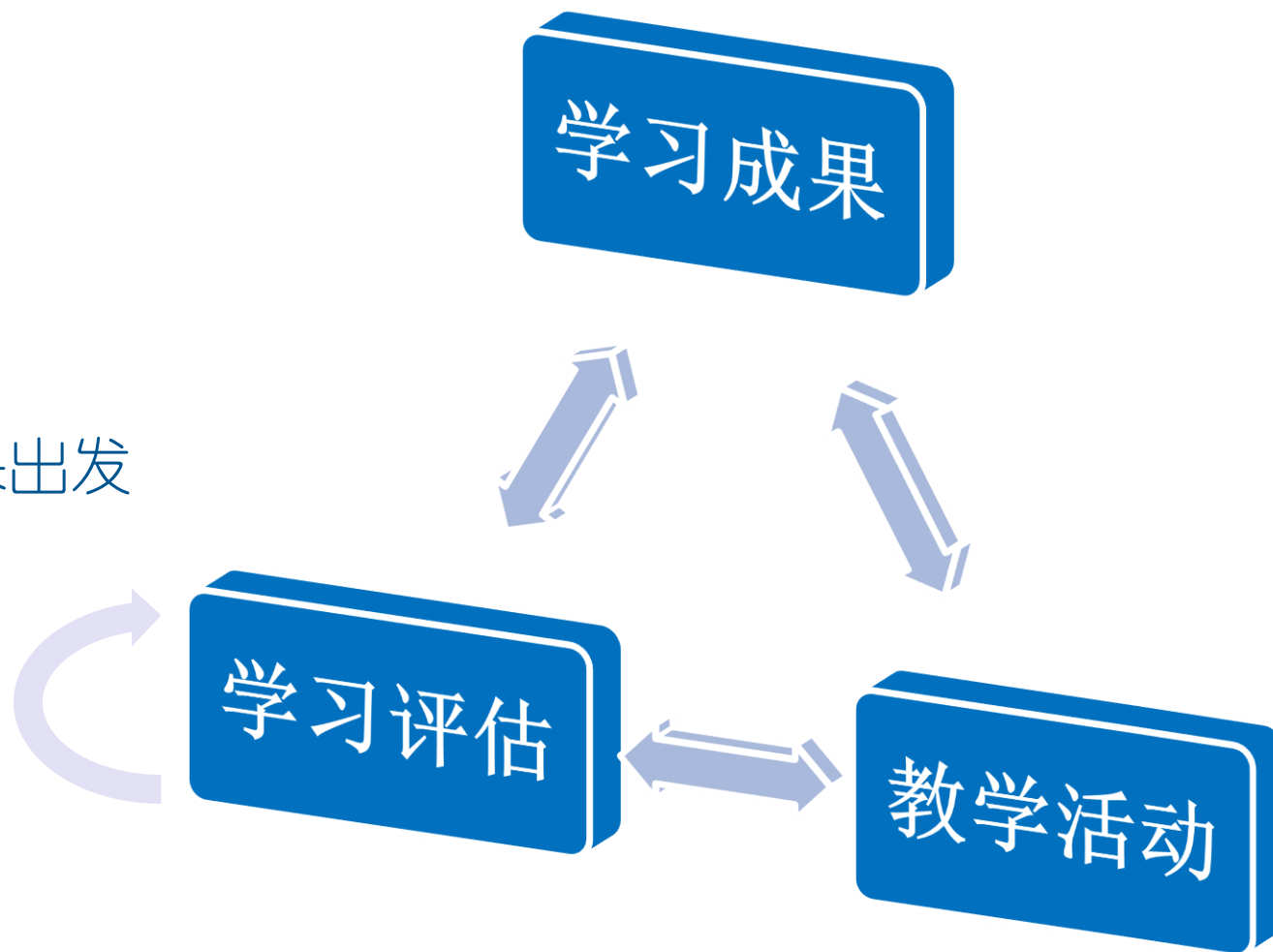
C. 写作

D. 行为观察

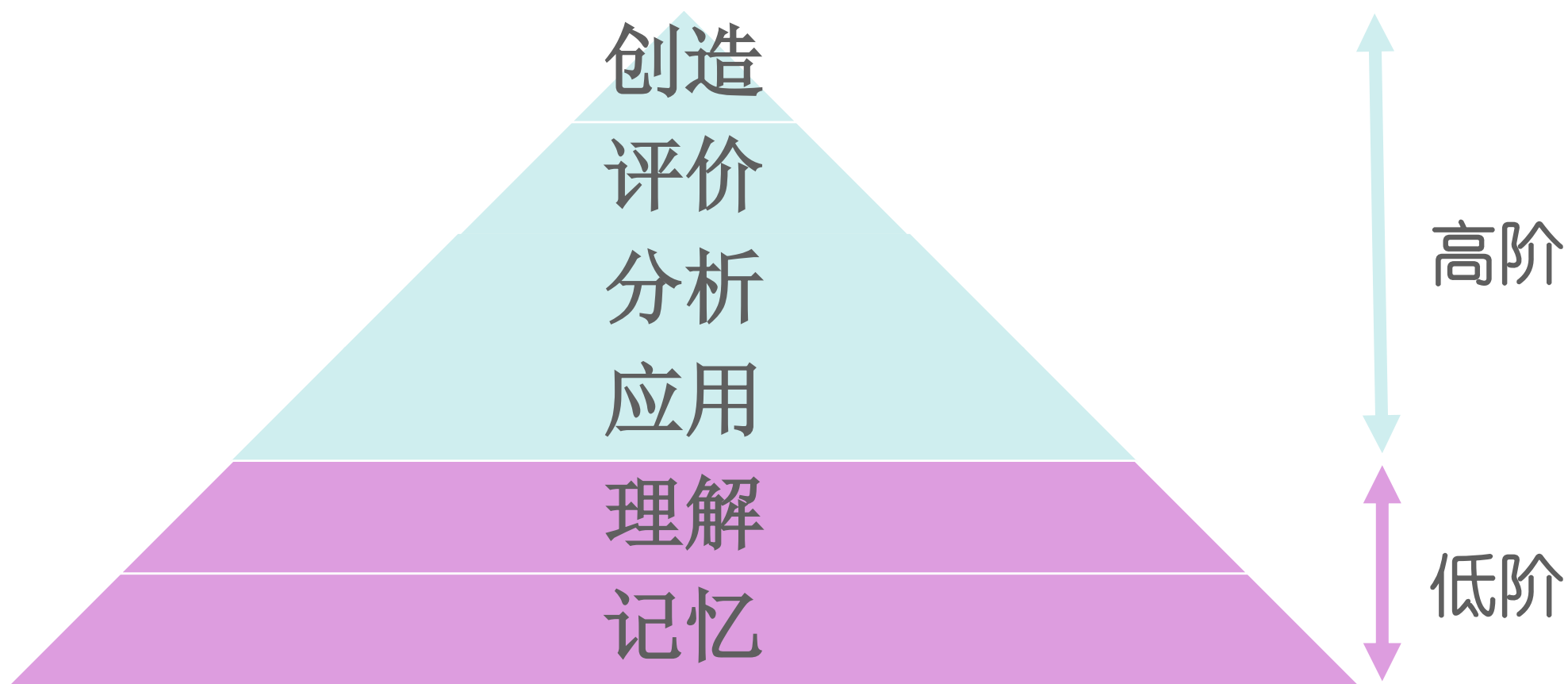
评估原则一

评估、教学活动对标学习成果

从明确、重要、可测量的学习成果出发



布鲁姆认知目标层次（修订版）



记忆

• 描述、复述、辨认、列举、匹配...

理解

• 转换、说明、讨论、解释、概括...

应用

• 计算、表现、采用、阐释、关联...

分析

• 分解、对比、批判、图解、论证...

评价

• 评估、推论、辩护、预测、鉴赏...

创造

• 合成、构建、设计、发展、发明...

评估原则二

公平、公正

✍ 任务难易度

✍ 评分标准一致性





2. 评估方法1（测试）



试卷规划（简化的双向细目表）

聚焦重要的学习成果
分配合适的比重

内容	记忆	理解	应用	分析	评价	创造
第一章	3	2				
第二章	4	3	1	1		
第三章	2	4		1		1
第四章	4	3	1			
第五章	3	3	1		1	
。 。 。						
题数	15	15	3	2	1	1

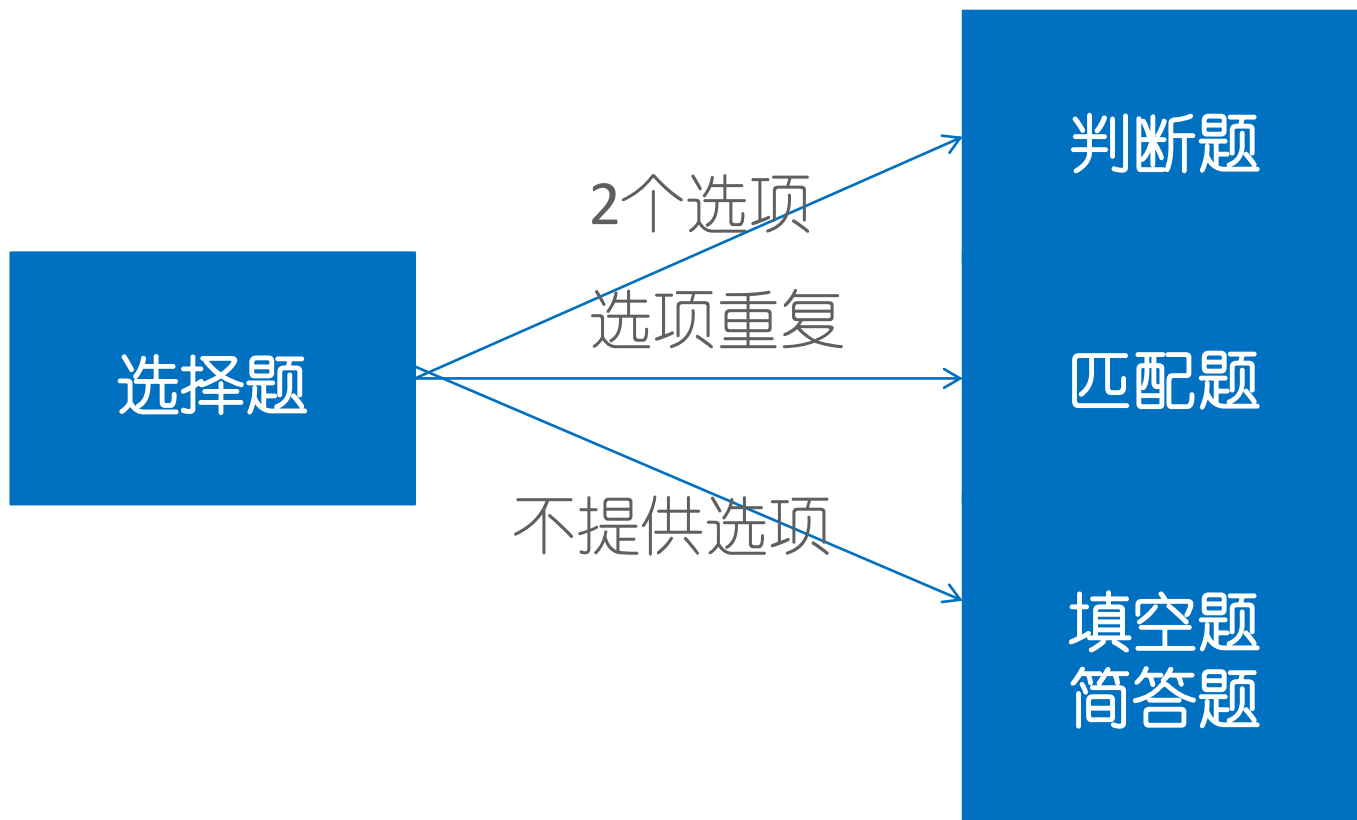
认知层次与题型



	选择*	是非	填空	简答	计算	作图	作文	设计
记忆	✓	✓	✓	✓				
理解	✓	✓	✓	✓	✓	✓		
应用	✓		✓	✓	✓	✓		
分析	✓						✓	✓
评价	✓						✓	✓
创造							✓	✓



选择题与其它题型



选择题的设计一例1

学习成果：解释细胞各成分的功能。

1a. 线粒体的主要功能是什么？

- A. 能量转换
 - B. 蛋白质合成
 - C. 蛋白质糖基化
 - D. 吸收营养物质
-

1b. 线粒体被称为细胞的“发电厂”，因为它们为细胞新陈代谢提供能量。

以下哪项观察最能支持线粒体的这个功能？

- A. 线粒体有双膜
- B. ATP发生在线粒体中
- C. 线粒体几乎存在于各种植物和动物细胞中
- D. 在线粒体中发现了克雷布斯循环的酶和末端呼吸所需的分子

选择题的设计一例2

学习成果：将先前获得的知识应用于给定情境。

2a. 记忆系统有哪些？

- A. 声象记忆
 - B. 短期记忆
 - C. 长期记忆
 - D. 单听记忆
-

2b. 钢琴调音师在其职业中主要使用以下哪个记忆系统？

- A. 声象记忆
- B. 短期记忆
- C. 长期记忆
- D. 单听记忆

选择题出题技巧—考查高阶认知能力



- ✍ 呈现一个场景（可以虚构），以下哪项原理适用？
- ✍ 呈现解决一个问题的几种方案，哪种是最佳方案？
- ✍ 呈现一个问题，可以使用本课中的哪个原理来解决它？
- ✍ 呈现一个实验，预测或解释实验结果。

在线考试案例—澳大利亚 University of New England

- ✍ 2017年开始实施有监考的在线考试
- ✍ 2019年54%的考试上线
- ✍ 2020预计100%的考试可上线

- ✍ 采用 ProctorU + Moodle
- ✍ 考试的窗口期一般为24小时，学生可以自由选择考试开始的时间
- ✍ 为了降低学生对于线上考试的技术焦虑，提供了Try-it exam
- ✍ 考试周2周，技术人员24小时值守，及时应对可能出现的技术问题。

在线测试的替代—口试



优势

- ✍ 能配合特定的学习目标—沟通能力、知识的理解和应用
- ✍ 即时的问与答可以避免误解
- ✍ 可以知道学生如何得到答案

不足

- ✍ 不擅长口头表达的学生，表现受影响
- ✍ 学生缺乏取舍权

口试一例1

- ✍ 致远学院 邢向军老师 《热力学与统计物理》，2015
- ✍ 学生 ~30人，
- ✍ 对重要概念口试，考查理解的深度



口试—例2：瑞典 KTH



Written vs. oral exam, teacher time

Written:

Constructing one exam and solution-sheet takes \approx 10-16 hours.

Correcting them takes \approx 20 minutes per student.

Oral:

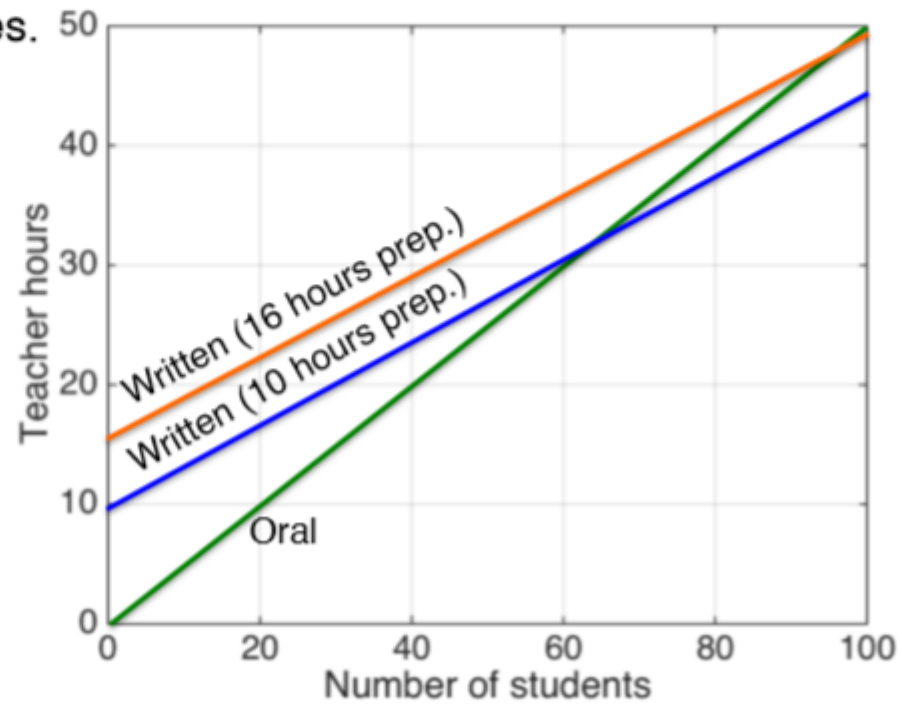
The exam takes \approx 30 minutes.

*But consider also
the re-exam!*

Let's see if we get it

- 16 hours to prepare exam
- 80% passing rate
- One re-exam

Break-even is at \approx 160 students



Kristina Edström,
The teaching
trick, 2019



3. 评估方法2（表现型评估）



表现型评估形式

- ✍ 口头表达任务（口头报告、辩论、角色扮演）
- ✍ 写作（研究小论文、主题论文、读书报告、案例分析、Wiki页面）
- ✍ 项目、设计（+ 展示）
- ✍ 异步讨论（讨论区讨论）



口头表达任务一例子：澳大利亚 Griffith University

✍️ 经济学专业，培养未来经济学家的口头交流能力



一年级，教家人、朋友一个经济学概念



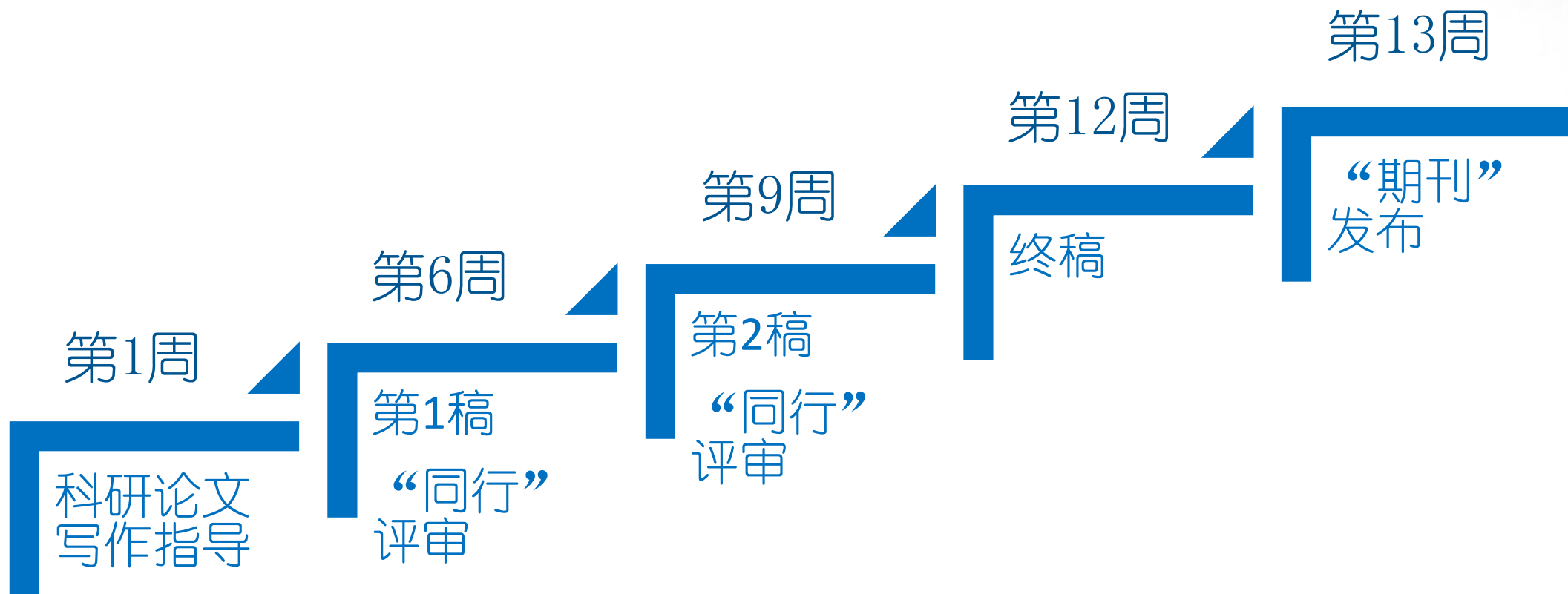
二年级，通过媒体与普通大众沟通宏观经济学理论



三年级，帮政府、私人企业沟通成本收益分析

写作案例 1—模拟期刊论文

《能源科学与技术》



Schulte, J. & Griffiths, N, 2014

写作案例 2

《微观经济学》

- ✍ 两篇500字以内的短文
- ✍ 用微观经济学原理，用普通人都能理解的话语，解释一个日常观察到的现象

“学习的经济学原理不仅不会在合上书本后忘得一干二净，反而会越记越牢”



Robert H Frank, Cornell University



写作案例 3



Prof. Rick Glofcheski
University of Hong Kong

《侵权法》

拍照及叙事

对媒体报道的与法律相关事件写日记



表现型任务的设计

- ✍ 是或类似真实职业领域的任务
- ✍ 是非良构 (ill-defined) 的问题 (没有现成解法)
- ✍ 需要持续的投入
- ✍ 需要寻找多种资源
- ✍ 最后是完整的作品





4. 表现型任务评估工具—量规 (rubric)



量规的作用

对教师

- ✎ 客观
- ✎ 效率
- ✎ 审视并调整学习目标及教学行为

对学生

- ✎ 引导有效学习
- ✎ 发展自我评估的能力

讨论区讨论参与量规

(By Frey., B.A.)

标准	不合格 0分	合格 1分	良好 2分	优秀 3分
频次	完全没有参与	当天参与1-2次	参与3-4次，但发帖未持续一周	一周参与4-5次
发帖	未发布讨论贴	发布了适量帖子，内容和准备较浅，没有涉及到要求的各个方面	所发帖子的内容涵盖了要求的各个方面，但没有对概念充分展开	所发帖子的内容涵盖了要求的各个方面且有具体展开
跟帖	没有回复他人的帖子	进行了浅层次的讨论（比如，表达了同意或不同意），未深入探讨	对他人帖子进行了详细评论	对他人帖子进行了分析，并且基于已有帖子拓展了有意义的讨论内容
内容贡献	帖子内容不正确、脱离主题、或与讨论无关	帖子内容重复，没有在讨论中提供实质信息	帖子内容符合事实，但缺少观点的全面发展	帖子内容符合事实，值得思考，对推动讨论有实质性贡献
参考与支撑	没有用参考或经验支撑	使用了个人经验，但没有参考阅读或研究	结合了一些参考文献和个人经验	使用了参考文献、阅读材料或个人经验来支撑观点
表达清晰度与技巧	帖子冗长、结构混乱或内容粗糙，有较多错误或内容不恰当	表达方式友善、礼貌且有帮助，但表达不够清楚或存在一些技巧问题	在讨论中提供了有价值的信息，表达较为清楚，但技巧上存在一些小题	评论清晰简洁，表达规范，易于阅读，没有语法和拼写错误

量规借鉴与改用



1. Rcampus: (量规集合)

<https://www.rcampus.com/indexrubric.cfm?nocache=1428304114104>

2. AACU VALUE rubrics: (通识教育)

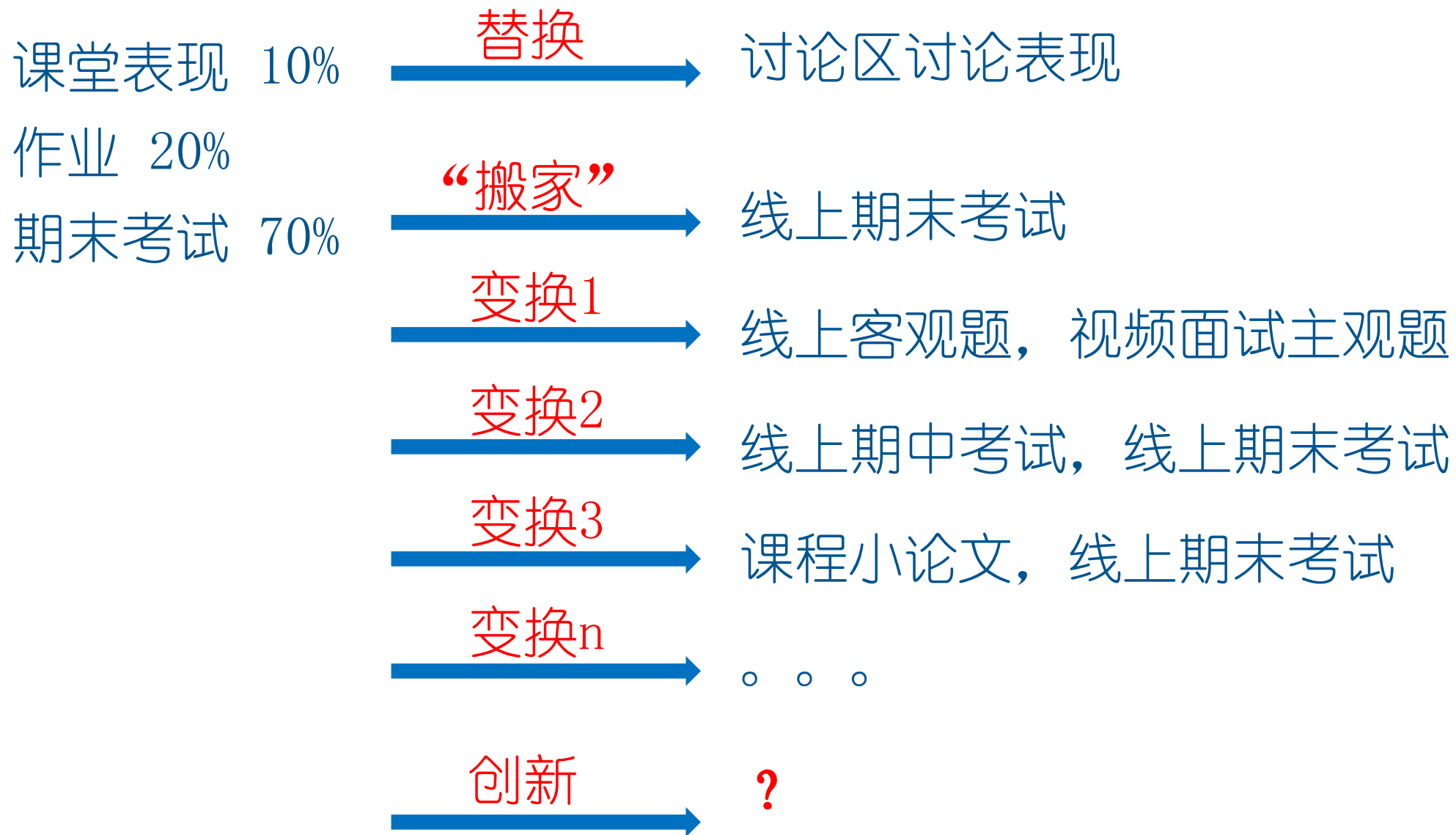
<http://www.aacu.org/value/rubrics/index.cfm>

个人可在课堂、课程网站、校内出版物上免费使用，声明：

"Reprinted [or Excerpted] with permission from Assessing Outcomes and Improving Achievement: Tips and tools for Using Rubrics, edited by Terrel L. Rhodes. Copyright 2010 by the Association of American Colleges and Universities."

3. Carnegie Mellon University rubrics: (论文、项目、口头展示、课堂参与)

<http://www.cmu.edu/teaching/design/teach/rubrics.html>



小结



评估目的和原则

- 总结性评估、形成性评估
- 评估对标学习成果
- 评估的公平公正

测试

- 试卷规划
- 试题设计技巧

表现型评估

- 任务设计原则

量规

- 对师生均有益



谢谢！

期待您的案例分享！