

# 哈佛大学“等级制”课堂评价模式的研究及启示\*

## ——以《大学物理》课程为例

张萍<sup>†</sup> 梁颖

(北京师范大学物理学系 北京 100875)

2015-05-22收到

<sup>†</sup> email: zhangping@bnu.edu.cn

DOI: 10.7693/wl20150707

**摘要** 课堂评价是教育教学活动中极为重要的一环,对教育和教学活动具有极强的导向作用。文章作者对美国哈佛大学的《大学物理》课程进行了一个学期的跟踪研究,深入分析该课程的课堂评价方式和具体操作办法,结合课程教、学、评等多视角探究其中蕴含的教育意义和对中国大学教育教学改革的启示。

**关键词** 课堂评价, 大学物理课程, 高等教育研究, 哈佛大学

中国的大学一直以来使用“百分制”评价学生,国际上许多大学则使用“等级制”。2015年5月15日清华大学宣布将以“等级制”代替“百分制”,引起众人热议。然而,大多数的讨论都停留在如何将传统意义的“百分制”换算成“等级制”和使用等级分制后如何进行成绩排名等操作层面上。事实上,“等级制”和传统“百分制”的评价模式在教育目标、教育理念和教学方法上都存在很大差异,只有真正理解“等级制”评价模式中蕴含的教育意义的本质,才能实现以课程评价模式的改革带动教育理念的更新和教育教学方法改革。

本文将详细介绍哈佛大学《大学物理》课程的“等级制”评价模式。通过对课程教学全过程的跟踪研究,包括一个学期的课堂观察以及对授课教授的访谈,从美国大学高等教育的目标、教学方法和教学评价三个方面,分析说明“等级制”评价中蕴含的教育学本质和意

义,为中国大学教育教学改革引入国际化视角,同时也为一线教师提供一个较为具体的教学案例,使教师们在教学实践中有章可循,便于理解、把握和运用。

## 1 引言

基于评价与学习的关系的不同,课堂评价分三种模式<sup>[1]</sup>:“关于学习的评价”(Assessment of Learning),“促进学习的评价”(Assessment for Learning),“作为学习的评价”(Assessment as Learning)。“关于学习的评价”又称终结性评价(Summative Assessment),发生在一个学习周期之末,如期中、期末考试、升学考试。它是对学习结果的评价,用来对学生划分等级和排名,其结果对学生的将来有直接的影响,所谓“一考定终生”,这是中国大学一直以来使用的“百分制”评价模式,这是一种重结果、轻过程的评价方法;“促进学习的

评价”又称形成性评价(Formative Assessment),发生在学习的全过程,评价的目的是多方面和及时地为教师和学生提供反馈信息,使他们不断调整教与学的步伐,共同努力实现教学目标;“作为学习的评价”其性质也是形成性评价,与促进学习的评价相比,更强调学生在学习过程中的自我评价和学生间的相互评价,并且把学生为主体的评价和他们的学习结合起来,认为评价即学习。

形成性评价是1967年美国大学教授斯克里文(Scriven)首次提到,许多研究者(Natriello, 1987; Fuchs, 1986; Crooks, 1988; Bangert-Drowns, Kulik, 1991; Kluger, DeNisi, 1996; Black, William, 1998)对课程评价也提出了一些原则,包含下列内容:

(1) 评价的反馈应该为学生清晰地描述学习目标上的进步和如何提高;

(2) 评价的反馈应该鼓励学生

\* 北京师范大学教育学部2015年度学科综合建议专项资金资助

进步;

(3) 评价在本质上应该是形成性的;

(4) 评价应该是经常性的, 贯穿于教学的始终。

美国在20世纪80, 90年代开始进行大规模教育教学方法改革, 传统的灌输式教学方法基本上被启发式、探究式、讨论式、参与式等教学方法取代, 目的是变“教”为“学”, 让学生主动地参与到学习中, 更多地强调自主学习和合作学习。教学方法的改革要求评价模式做出相应的变革, 教育者开始使用形成性评价来实现对多元化学习过程的评价, 其中包括学生的学习兴趣、态度和努力程度以及自主、合作和创新等能力……形成性评价贯穿于整个教学过程, 目的是引领教师和学生有效地教和学、促进每个学生的成长和最大化的发展。教育教学改革中, 教育者意识到评价每个学生“是否尽了全力”比评价他们中“谁最优秀”更加重要, 此时更具有选拔意义的“百分制”评价

模式就失去其价值, 逐渐被“等级制”评价模式所取代。下面以哈佛大学《大学物理》课堂评价为案例, 详细分析“等级制”评价中蕴含的教育意义。

## 2 哈佛大学《大学物理》课程的学业成绩构成

在哈佛大学, 教师可以依据课程设计, 自主确定课程的评测方法, 但必须写入课程信息并在课程开始前告知学生。本文介绍的是物理11(Physics 11)的评价模式, 这门课程与中国《大学物理》课程对应, 是一门面向全校理工科学生开设的基础课程。中国大多数高校物理实验课程是单独设课、单独评价, 哈佛大学物理实验是配合理论课开设, 作为同一门课程整体评价。

我们先介绍这门课程的分数组成, 然后再结合课程详细介绍具体操作办法。这门课程的学业成绩中包含了课前、课中、课后、实验和测验与考试多个评价环节, 分数构

成如表1。

在以百分制计算得到学生总分后, 需要根据分数所在的区间折算成等级分数(Grade), 折算方法见表2。哈佛大学学生课程成绩使用等级评定, 以ABCD形式记录。

## 3 课程使用的教学方法与课堂评价模式

为了深入理解哈佛大学物理课程考试和评价中蕴含的教育意义, 我们先详细介绍一下这门课程各教学环节的教学目标、使用的教学方法、评价方式和评分细则。

### 3.1 课前阅读——即时反馈法 JITT(Just-in-Time Teaching)

JITT<sup>[2]</sup>是一种基于网络的互动教学策略, 它要求学生在上课前阅读课程相关材料并在网上回答课程相关问题, 教师根据学生回答问题的情况, 及时得到他们学习成效的反馈, 调整课堂教学的内容和教学策略。哈佛大学物理课使用了这一教学方法: 教师给出“阅读任务单”指导学生阅读新课相关内容, 学生在网上提交阅读作业, 网络在上课的前一天晚上关闭。在对阅读作业的评价中, 重点评价学生是否努力学习, 不评价答题的对错, 一学期中学生少于三次没有提交阅读作业是不扣分的。

### 3.2 课上讨论——同伴教学法 (Peer Instruction)

哈佛大学物理课程课堂教学使用同伴教学法<sup>[3]</sup>, 将一节课分成几个单元, 每个单元围绕一个核心概念, 以概念测试题(Concept Test)组织教学。在课程教学中基于概念测试题, 利用教室交互系统CRS(Classroom Response System)先让学生发送个人答案, 然后与邻座同学

表1 哈佛大学《大学物理》课程分数构成

评价项目	分数比例
课前阅读	10%
课上讨论	6%
课后作业	15%
物理实验	15%
期末考试	18%
在下面三个成绩中取最好的两个:	
第一次阶段考试	18%
第二次阶段考试	18%
期末考试	18%
合计	100%

表2 课程分数等级与最低分数(百分制)对应表

等级	A	A-	B+	B	B-	C+	C	C-	D+	D	D-
最低分数	90	84	78	73	68	64	60	56	52	48	44

进行讨论，之后再次发送答案。CRS将汇总学生答案的分布以柱形图形式显示在教室的大屏幕上，教师和学生可以立即得到反馈信息。CRS事先已经注册了学生的个人信息，在学生回答问题时，学生出勤与否和参与课堂活动时的部分信息都被CRS记录下来。哈佛大学物理课程对学生课堂讨论的评价标准是：只评价学生是否积极参与课堂讨论，不对学生答题正确与否评分，在整个学期的课堂活动记录中，学生缺少三次以下不被扣分。

### 3.3 实验——团队工作

在哈佛大学课程设置中，物理实验配合物理理论课开设，作为一门课程统一评价。由负责实验的教授依据理论课的内容设置实验题目，一学期中共有5个实验，每个实验3分，共15分。实验是3人一组，要求学生预习，每个学生在实验之前须独立回答和实验有关的问题(pre-lab questions)用来检查学生预习和实验准备情况，占1分，实验中学生个人表现，占1分。一个实验小组共同完成一份实验报告，获得团队分数，占1分(团队分数是合作学习的一种评分模式，它只对小组合作的最后成果评定分数，不

评价个人的贡献大小)。

### 3.4 课后作业——独立完成之后小组讨论

学生需要完成每周一次的课后作业。这门课程一周有两次课，分别在星期二和星期四。课后作业于星期四的课上布置，要求在下一周的星期四下午5点以前交到指定的信箱中。最晚可以在接下来一周的星期一上午10点之前交，晚交的作业只能得到一半的分数。要求学生必须先独立完成作业后，带着其中的问题和自己的观点参加小组讨论(在课程开始时学生成立学习小组)，再对自己的作业进行评价并做出改进，上交的作业中必须包含个人原始的解决方案和小组讨论后改进的方案，提交的个人作业上必须注明讨论小组成员的名字。课后作业的评价模式是个人成绩，所以和别人作业极其相似的作业是不被接受的。

### 3.5 阶段测验与期末考试——开卷考试并可以选择考试内容

哈佛大学的这门物理课包含两次阶段测验和期末考试。阶段测验只涉及该学习阶段的课程内容，期末考试包含整个学期的课程内容。三次考试都允许看书和笔记(open book, open notes)。答题必须有完

整的过程和合理的解释，否则会被扣分。如果只有正确答案，完全没有合理的解题过程，将得不到任何分数。考试时允许学生选择考试内容，两次阶段考试的试卷中都包含四部分内容，学生只需要选择完成其中的三部分。期末考试有八部分内容，要求选择完成其中的六个部分。试卷中的部分题目有奖励分数(bonus)，鼓励学生用多种方法解决问题或对问题给出深度的解释。

在合成课程总分数时，期末考试分数占18%；在两次阶段考试和期末考试三个成绩中选取最好的两个成绩各占18%。按照这个规则，如果期末考试成绩在三次考试中不是最差的，在总成绩中将被计算两次，以取代阶段考试中那个低的分数。图1是这门课程两次阶段考试和期末考试的学生成绩分布图<sup>[4]</sup>，注册课程的学生可以在网上看到自己的成绩和全班学生的成绩分布，知道自己的成绩在全班学生中的大致位置。

## 4 “等级制”评价模式中的教育学意义和启示

中国大学的学业成绩评价模式

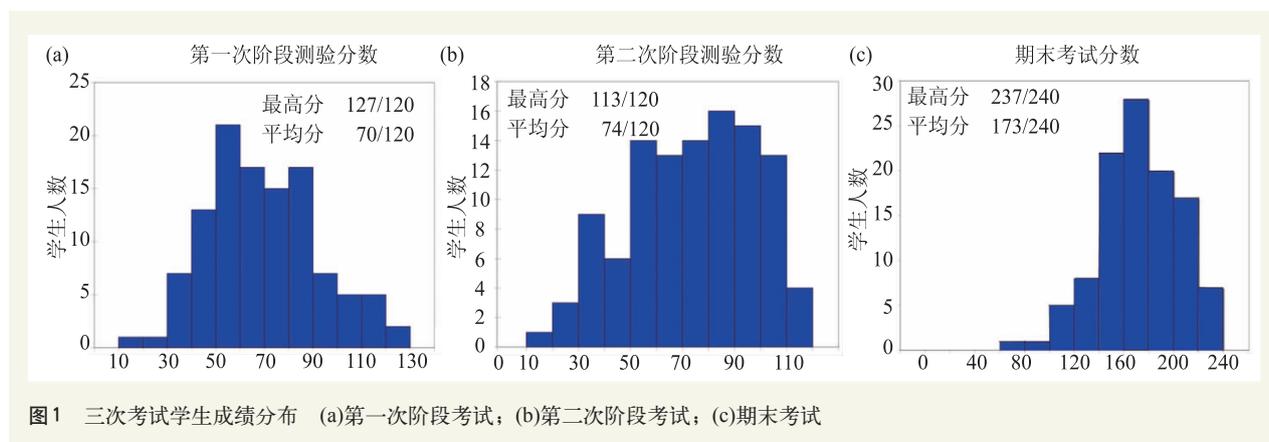


图1 三次考试学生成绩分布 (a)第一次阶段考试；(b)第二次阶段考试；(c)期末考试

一般都使用“百分制”，评价时比较侧重于知识内容的完整和给予学生一个精确的评分，在课程教学中重点关注课程(Curriculum)目标，即掌握学科知识，学生在大学课程中学到很多的知识，但很少获得实际练习和发展技能的机会。美国大学在课程教学中则更加强教育(education)的目的，哈佛大学著名哲学家和教育学家伊斯雷尔·谢弗(Israel Scheffler)说：“教育的目标必须包含判断习惯的养成、品格的培养、标准的提高、促进理解、品位提高与能够区分善恶、充满好奇与善于质疑、追求时尚和提高美感、渴求新理念和探求未知<sup>[5]</sup>。”这些教育理念很好地体现在哈佛大学课程教学和课堂评价之中。

#### 4.1 强调学生自主学习

“为什么让学生自主学习？”作者访谈了哈佛大学两位讲授物理课程的教授。埃里克·马祖尔(Eric Mazur)说：“在课堂上教师简单地重复教材上印刷的内容，意味着教师认为学生不知道如何阅读教材，是对学生自学能力的轻视。”另一位教授加里·费尔德曼(Gary Feldman)也表达了同样的观点，他说“教师在课上讲那些学生可以自学的内容是不负责任的，因为这样做的结果是在钝化学生的自学能力。”目前哈佛大学的课堂上教师讲的很少，大多数时间是要求学生讨论，为了课堂的讨论能有效进行，所有课程中无论是文科还是理科，都要求学生在课前大量阅读，准备充分了，上课时才能贡献自己的想法，很好地与别人交流，否则无法有效地参与到课堂讨论中。哈佛大学非常重视学生的主动学习和自主学习。为了引导

学生有效地自学和让教师可以针对学生自学后的水平组织课堂教学，必须及时对学生的自学成果进行评价和反馈。

强调自主学习的重要意义在于，学生必须对自己的学习负责任，教师则在启发和支持学生自主学习时发挥重要作用。学生在自主学习的过程中，逐步形成自己的学习方法，反思和调试学习过程，监测学习进展，自我评估，从而学会对自己的学习负责任。研究表明，自主学习不仅可以提高学习动机、学业成绩和自信心，也有利于学生认知技能、元认知技能和情感技能的发展，还将使学生获得终生学习能力和职业发展上的高成就<sup>[6]</sup>。

#### 4.2 构建的多样化的合作学习环境

美国哈佛大学文理学院前院长亨利·罗索夫斯基(Henry Rosovsky)曾经自豪地说过这样一段话：“我们哈佛大学学生从互相之间学到的知识比从老师那里学到的更多。”从哈佛大学的物理课程评价模式中可以看出，合作学习渗透在课堂、实验室和课外的家庭作业等多个环节，对合作学习的评价方式也是多样化的。

为了鼓励学生在课堂上积极参与讨论，有效地相互合作，在课堂学生小组讨论的评价中，不评价学生观点的正确与否，只评价学生是否积极地参与讨论，学会思考比给出正确答案更重要。

哈佛大学物理实验课程的目标除了培养实验技能外，还有一个重要的目标是培养学生团队合作的能力。在评价分数构成上包含个人分数和团队分数。个人分数，评价每

个学生是否做好实验前的各种准备工作和实验过程中表现出的各种实验能力。实验完成后，要求整个小组对实验结果进行讨论、分析并得出结论，如果不同的小组成员的观点有分歧，则必须达成共识后，大家共同完成一篇实验报告，教师对这份报告评分，同一个实验小组的成员无论个人能力高低、贡献大小都得到相同的小组分数，从而实现在合作学习中对个人责任和团队合作两个要素的评价。

课后作业要求学生必须在独自完成作业后，参加4个小时左右的小组合作学习，围绕家庭作业展开讨论，各自说明自己的解题思路，为学生提供机会相互争辩、讨论、交流，共同解决问题，对课程内容形成更丰富、更深刻的理解。依据小组讨论的结果，学生可以对自己的作业进行自我评价和改进，作业的评价模式是提交个人作业获得个人成绩，作业中必须包含个人原始的解决方案和小组讨论后改进的方案。抄袭作业被认为是学术不端。

研究结果一致表明，与竞争化学习或个人努力相比，合作学习可以产生以下效果：对学科领域和学习更积极的态度；对学习的更大成就动机和内在动机，更多高层次的推理，更深入的理解和批判性思考，更能从他人的角度看待问题，更强的社交能力等<sup>[7]</sup>。

#### 4.3 获得学科能力比全面记忆学科知识更重要

中国大学物理课程的考试大多是闭卷的形式，学生在学习和备考过程中，为了取得好成绩需要记忆大量的公式和题型。与此不

同, 哈佛大学物理课程的阶段测验和期末考试都允许学生看书和笔记, 学生无需记忆公式和题型, 考察的重点是, 学生是否真正理解课程知识和能否灵活地应用所学知识解决新问题。美国大学的学业评测很少考察记忆性、事实性知识, 重视较为复杂的学习成果, 比如理解、应用、分析、评价和创造。

在课程学习中, 有些学生按部就班, 照本宣科地完成所有的课程内容。有些学生则有强烈的好奇心和丰富的想象力, 在学习过程中对教科书和教师讲授的内容不是简单接受, 而是冥思苦想、提出问题、深入钻研, 超出教科书查阅资料甚至链接多学科知识, 深刻理解和灵活运用所学内容, 这样做会花费大量时间和精力, 有些学生可能因此没有时间很好地学完课程的全部内容。两类学生中哪一类学生应该得到鼓励而获得好的成绩呢? 哈佛大学选择鼓励学生深度学习, 他们在考试内容上给予学生选择空间, 学生无须完成试卷上的所有题目, 而是根据自己的学习情况和特长, 按照考试要求选定自己具有优势的题目。在试卷的开始有这样一段提示: “最好的选择是出色地完成选定的题目, 而不是做完所有题目”。某些试题是有附加分数的, 如果学生能够有独特的思维方式, 用多种方法、多角度解决和诠释问题就可以获得额外的奖励分数, 并且上不封顶, 如图 1(a)所示, 第一次阶段测试中满分 120 分, 最高分 127 分。

哈佛大学的这些举措意在鼓励学生更多高层次的推理, 更深入的理解和批判性思考、渴求新

理念和探求未知。充分体现了在课程教学中培养创新人才的教育理念。

#### 4.4 鼓励学生进步, 促进所有学生全面发展

中国大学的物理课程采用常模参照评测方式(选拔性评价), 使用“百分制”来对学生排名次, 比高低, 以此甄别和选拔“优秀人才”, 获得保送研究生和出国深造等资格, 学生在充满竞争的环境中学习。为了“公平”选拔, 只有采用统一考试, 包括统一的试题内容和评分标准, 很多大学都要求分数分布符合正态分布。

相比之下, 哈佛大学物理课程的评价模式显得不是很“公平”, 具体表现在四个方面: 一、学生可以选择考试内容, 不同学生的考试内容是不同的; 二、合成课程总分数时, 如果期末考试成绩超过两次阶段考试中的任何一次, 在总成绩中将被计算两次, 以取代阶段考试中那个低的分数; 三、物理实验中的团队分数(team mark), 只评价小组合作的最终成果, 小组的每个成员无论个人表现如何都获得同样的分数; 四、不要求分数的正态分布, 如图 1 所示。课程成绩合成使用百分制, 成绩记录使用等级评定(ABCD), 失去了“百分制”分数给学生排定名次的“精度”, 其实, 由于前面三个方面的“不公平”, 哈佛大学课程评价过程中百分制分数本身就不具有传统意义的“精度”。

表面看来“不公平”的四个方面都蕴含着深刻的教育意义: 一、给学生更多的选择, 充分发挥他们的兴趣与特长; 二、大学教学的目标不是对学生进行筛选和分类, 而是激励学生不懈努力和全面发展;

三、课堂教学应该为现实世界的工作做准备, 学生需要在大学的学习中, 学会如何通过有效合作, 使个人和团队的成就最大化; 四、美国大学中一般采用标准参照评测, 使用等级评价, 目的不是为了选拔学生而是评价学生在多大程度上达到了预期的学习成果, 并且奖赏不受限制。这样的评价模式可以鼓励学生为了实现共同的目标, 一起学习, 相互支持, 共同进步。大量对比研究清楚地表明, 个人成就和竞争成负相关, 与之相反, 学生间积极的相互依赖关系能促进彼此的学习成绩提高, 使每个人的学习成果最大化, 从而使更多的学生获得高成就<sup>[8]</sup>。奖赏不受限制并不意味着标准降低, 如图 1 所见, 这门课程两次测试和期末考试的平均分为: 70/120, 74/120, 173/240, 折算成百分制是: 58, 61, 72。这个案例中有 20% 的学生总成绩是不及格的。在哈佛大学, 学生必须要拼尽全力地学习, 如果还是不及格, 说明自己不适合学习目前的专业, 应该考虑尽早地转学其他的专业和学校, 哈佛大学每年有大约 20% 的学生会因为考试不及格或者修不满学分而离开原来的专业或学校。相反, 在中国, 我们尽管使用选拔性评价, 但是大学严进宽出, 学生在竞争的环境中学习却很少有淘汰。对于多数学生来说, 即使自己再努力学习也超不过“学霸”, 而不用功学习也不会“挂科”, 于是, 这些学生在大学期间就可以享受到四年他们童年时候失去的美好时光。

## 5 结束语

在这个全球化的时代, 引进、

借鉴发达国家的教育教学理论和实践经验是促进中国教育教学发展的重要途径和有效策略。然而,我国大学传统意义的“百分制”和目前美国大学常用“等级制”之间的差异,不只是学生学业成绩呈现方式的不同,其中蕴含两者在教育理念、教育目标、教学方法和对学生的评价方式等多方面的巨大差异。要想有效地将“等级制”评价模式落地到中国大学,使中国的高等教育和世界接轨,需要对中国大学教育进行一次全方位的改革。

首先我们必须厘清,大学希望达到怎样的目标,大学应该让

学生通过四年的学习收获什么?在人生的关键时期,大学应该怎样帮助学生成长与发展?哈佛大学前校长德雷克·博克(Derek Bok)曾列出适合所有本科生的一些目标<sup>[9]</sup>,它们是:表达能力、批判性思维能力、道德推理能力、公民意识、适应多元文化的素养、全球化素养、广泛的兴趣、为就业做准备。他还说:“也许有些人认为我遗漏了一些有价值的目标,如培养学生的想象力、创造力、领导力等。”

接下来的要解决的问题是,在大学中许多教师认同很多有价值的目标,却不懂得怎样去实现

这些目标。很多教育研究都表明,改革大学课程和教学内容对实现上述目标作用不大,改革教学模式和教学方法才是最有效的办法,然而每次到了大学的课程改革时,大家讨论最多的还是应该开设哪些必修课和选修课。因此,要想有效实现大学的目标,还必须进行大规模的教学方法改革。只有当教师理解大学教学目标,并使用有效的教学方法实现这些目标,才能真正明白应当如何评价每位学生,有效地促进他们成长和实现他们的最大化发展。

## 参考文献

- [1] Noyce P E, Hickey D T. *New Frontiers in Formative Assessment*. Cambridge, MA: Harvard Education Press, 2011
- [2] Gregor M, Novak, Evelyn T *et al.* *Just in time teaching*. Upper Saddle River, NJ: Prentice-Hall, 1999
- [3] 张萍, Mazur E. *中国大学教学*, 2010, (08):69
- [4] <https://www.physics.harvard.edu/11b>
- [5] Scheffler I. *Teachers College Record*, 1976, 78(2):206
- [6] Cleary T J, Zimmerman B J. *Psychology in the Schools*, 2004, 41(5):537
- [7] Johnson D W, Johnson R, Smith K. *Active learning: Cooperation in the college classroom*. Edina, MN: Interaction Book Company, 1998
- [8] Johnson D W, Johnson R. *Cooperation and competition: Theory and research*. Edina, MN: Interaction Book Company, 1989
- [9] Bok E. *Our Underachieving Colleges: A Candid Look at How Much Students Learn and Why They should Be Learning More*. Princeton, NJ: Princeton University Press, 2008

## 读者和编者

## 《物理》有奖征集封面素材

均有稿酬及全年《物理》杂志相送。

请将封面素材以附件形式发至: physics@iphy.ac.cn; 联系电话: 010-82649470; 82649029

期待您的参与!

为充分体现物理科学的独特之美,本刊编辑部欢迎广大读者和作者踊跃投寄与物理学相关的封面素材。封面素材要求图片清晰,色泽饱满,富有较强的视觉冲击力和很好的物理科学内涵。被选用的封面素材提供者,

《物理》编辑部