

高中骨干教师命题能力培训课程（一）

# 现代教育测量与考试评价的 理论及方法

王后雄

华中师范大学考试研究院  
院长·博士·教授·博士生导师



# 目录

## CONTENTS

一

教育测量与考试评价的理论基础

二

教育考试题型建构的理论与方法

三

构建中国特色教育考试评价体系

四

高考 / 学考试题难度的控制技术

五

高考 / 学考双向细目表创新设计

六

教育考试的命题流程及评价标准

七

高考 / 学考评分标准规则的制定

八

高中学考试卷评价指标体系设计

九

“素养为本”试题分析模式示例

十

开展教育考试与评价的实践研究



# 教育测量与考试评价的理论基础

## (一) 心理学理论

心理学理论是教育测量理论的基础，表1所示是最具代表性理论。

表1 心理学理论及其贡献

代表人物	依据	成果
布卢姆 (Bloom, B.S)	认知心理学	心理认知过程分为记忆、理解、运用、分析、综合和评价6层次。
加涅 (Gagne, R.M)	根据学习的复杂程度	将学习划分为八个层次，被称为“学习层次理论”，即信号学习、刺激反应学习、链索学习、言语联想学习、辨别学习、概念学习、原理或规则的学习和解决问题的学习。
	学习的本质	五种学习结果：智慧技能、认知策略、言语信息、态度和动作技能方面的改善。
梅耶 (Mayer, R.E)	学习过程	学习过程分为无学习、机械学习和有意义学习三类，提出认知过程可分为记忆、理解、应用、分析和评价。
王磊、冯忠良、郑国民等	学科基本能力	提出学习理解（辨识记忆、概括关键、说明论证）、应用实践（分析解释推理预测、简单设计）、迁移创新（复杂推理、系统探究、创新思维）3×3学科能力要素。

心理学相关理论证明，人的认知过程具有层次性、顺序性和系统性，这为知识、能力、素养层次水平框架的制订提供了参考和支持。

## (二) 教育测量学指标

### 1. 试卷的信度 (r)

信度就是试卷的可靠性，它是反映测量一致性程度的指标。一般要求0.90以上。信息系数 ( $Y_{xx}$ ) 是一组测量分数的方差 ( $S_t^2$ ) 与观测分数方差 ( $S_x^2$ ) 的比率。

$$Y_{xx} = S_t^2 / S_x^2$$

在实际测量中，信度不能由公式求得，而只能根据一组实得分数作出估计。一般来说，测验题量越多（长度），信息值越高。

### 2. 试卷的效度 ( $\alpha$ )

效度是指测量的有效性，即一个测验对它所要测量的特质准确测量的程度。

$$\alpha = S_v^2 / S_x^2$$

$S_v^2$ —与测量目的有关的变异， $S_x^2$ —实测值变异。

可信的测验未必有效，而有效测验必定可信。信度是效度的必要条件。

### 3.试卷（题）的难度（P）

$$P = \frac{\bar{X}}{a}$$

（ $\bar{X}$ 为全体考生实得的平均分， $a$ 为满分值）

高考：0.50-0.65，选考：0.60-0.70，学考：0.80-0.85。

### 4.试题的区分度（D）

$$D = P_H - P_L$$

（ $P_H$ —高分组通过的人数比例， $P_L$ —低分组通过的人数比例）

$D > 0.40$ ，非常优良， $D < 0.20$ ，必须淘汰或改进。

一般来说，较难、较易的试题区分度低，难度系数接近0.50时，区分度最大。

### （三）项目反应理论

一组具有共同基本思想的试题应答模型或项目反应理论模型的统称。其基本思想是：考生对某道试题的应答可以看做是考生的特征和试题特征的函数。

**优点：**用试题和考试的信息函数来表示考生能力水平估计值的测量精度。

**说明：**应用项目反应理论需要专门计算机软件、测量专家与考试开发者共同完成。

#### (四) SOLO分类理论—学习质量的评价

澳大利亚著名教育心理学教授（John Biggs, Collis, K.1982）提出，一种可观察到的学习质量的结果，在教育测量学领域产生了广泛影响。

SoLo分类理论对个体的学习结果进行分类，将学生的认知发展分为五种类型，其基本含义可用图1表示：

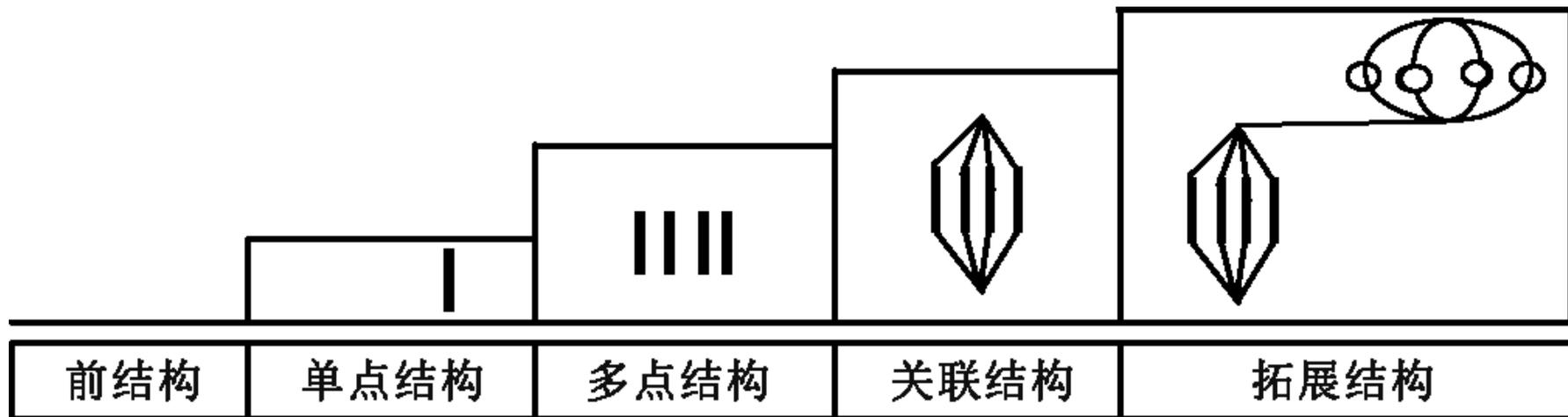


图1 SOLO学习结果分类

**1.前结构水平：**学生对学习任务本身不能进行恰当的处理，只是重复问题，还不能解释问题。

**2.单点结构水平：**能提出与学习任务有关的一个方面，但在诸要素之间或思维之间联系，只能理解问题的很少一部分。

**3.多点结构水平：**能提出与学习有关的几个独立方面，或能理解任务的许多方面，但相互之间无法形成关联。

**4.关联结构水平：**能把相关的方面整合成一个前后一致的整体结构。

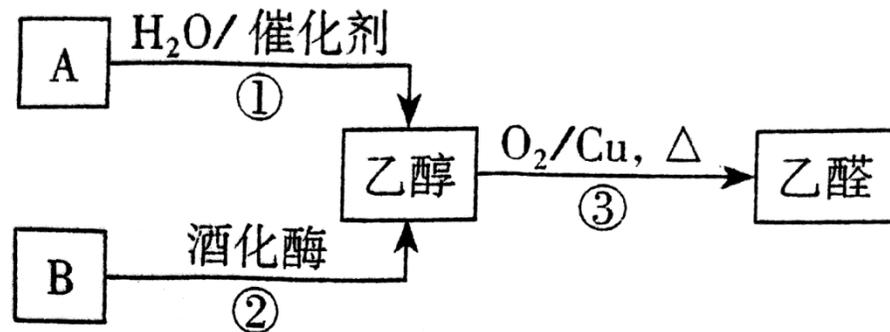
**5.拓展抽象水平：**能把连贯的整体概括为或再概念化为一个抽象的更高水平。

随着学者对SOLO理论深入研究，可以细分形成9个水平。

现以浙江省2015年10月起的三次学考/选考化学科目考试26题为例，用SOLO理论进行学习成果结构分析。

[2015年10月浙江省普通高校招生学考/选考科目考试，以下简称1510题] 26.  
(4分)

气态烃A中碳元素与氢元素的质量比为6：1。淀粉在一定条件下水解可生成B，B在人体组织中发生缓慢氧化，放出热量，提供生命活动所需能量。有关物质的转化关系如下图：



请回答：

(1) ①的反应类型是\_\_\_\_\_。

(2) 向试管中加入2 mL 10%氢氧化钠溶液，滴加4~5滴5%硫酸铜溶液，振荡后加入2 mL 10% B溶液，加热。写出加热后观察到的现象\_\_\_\_\_。

(3) 反应③的化学方程式是\_\_\_\_\_。

[2016年4月浙江省普通高校招生学考/选考科目考试, 以下简称1604题]

26. (6分) A是一种重要的化工原料, 部分性质及转化关系如下图:

请回答:

(1) D中官能团的名称是\_\_\_\_\_。

(2) A→B的反应类型是\_\_\_\_\_。

A. 取代反应

B. 加成反应

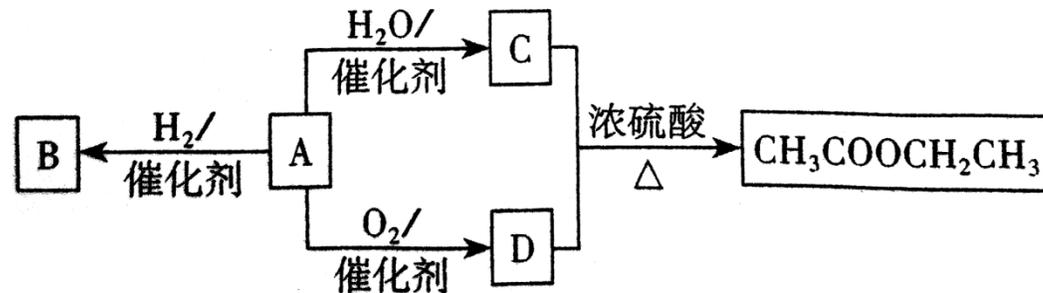
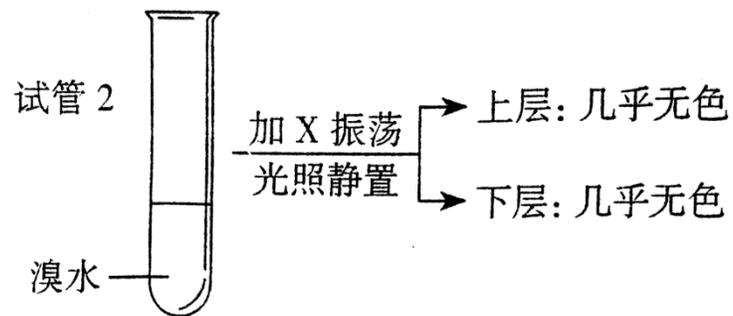
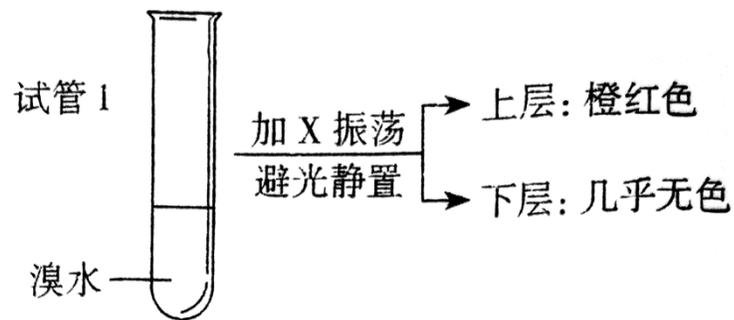
C. 氧化反应

D. 还原反应

(3) 写出A→C反应的化学方程式\_\_\_\_\_。

(4) 某烃X与B是同系物, 分子中碳与氢的质量比为36:7, 化学性质与甲烷相似。现取两支

试管, 分别加入适量溴水, 实验操作及现象如下:



有关X的说法正确的是\_\_\_\_\_。

A. 相同条件下, X的密度比水小

B. X的同分异构体共有6种

C. X能与Br<sub>2</sub>发生加成反应使溴水褪色

D. 试验后试管2中的有机层是上层还是下层, 可通过加水确定

[2016年10月浙江省普通高校招生学考/选考科目考试, 以下简称1610题] 26.

(6分)

烃A是一种重要的化工原料。已知A在标准状况下的密度为 $1.25 \text{ g}\cdot\text{L}^{-1}$ , B可发生银镜反应。它们之间的转化关系如图:

请回答:

(1) 有机物B中含有的官能团名称是\_\_\_\_\_。

(2) 第⑤步发生加成反应, 该反应的化学方程式是\_\_\_\_\_。

(3) 将金属钠与有机物D反应所得的少量产物溶于水, 滴加2滴酚酞试液, 水溶液显红色, 其原因是\_\_\_\_\_ (用化学方程式表示)。

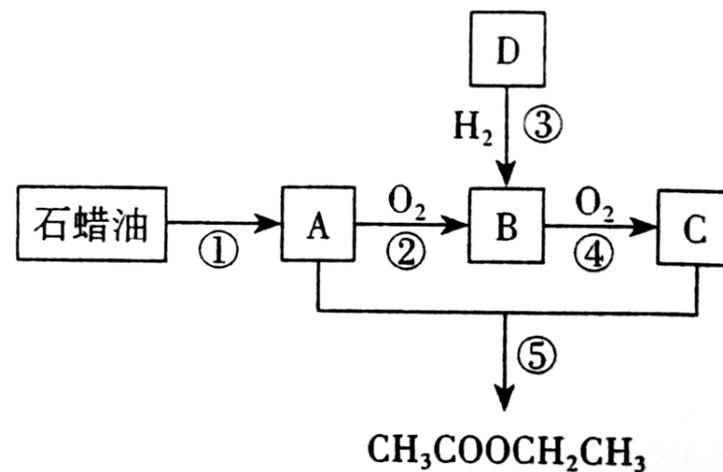
(4) 下列说法正确的是\_\_\_\_\_。

A. 有机物A与D在一定条件下可反应生成乙醚  $[(\text{CH}_3\text{CH}_2)_2\text{O}]$

B. 用新制碱性氢氧化铜悬浊液无法区分有机物B、C、D

C. 通过直接蒸馏乙酸乙酯和有机物C的混合物, 可分离得到纯的乙酸乙酯

D. 相同条件下乙酸乙酯在氢氧化钠溶液中水解较在稀硫酸中更完全



可观察的学习成果结果是约翰·彼格斯和凯文·科利斯教授以皮亚杰的发展阶段论为基础建立起来的。认为学习者回答某个问题时所表现出来的思维结构是可以检测的，可以判断处于哪一层次。该理论目前在国际上已被广泛应用于各学科的考评上。以可观察的学习成果结构评价理论为基础，分析3份试卷的第26题，可以更好地认识各小题设问所处的能力水平层次，明确必考阶段有机化学教学所要到达的教学目标，见表2所示（沈明祥，2017）。

表2 从可观察的学习成果结构评价理论（SOLO）详解第26题

		1510题	1604题	1610题
学习 成果 结果	前结构水平*			
	单一结构水平	1. **、 (3)		(4) D
	多元结构水平	(2)	(2)、(3)	
	关联结构水平		(1)、(4) ACD	(1)、(3)、(4) BC
	拓展抽象结构水平		(4) B	(2)、(4) A
参加考试的学生选考、学考 人数概况***		11.3万人参加 学考, 约1800 人参加选考	具体不详	选考考生约为12万人, 占化学科总考生数 45.5%
平均 分	未除去零分的平均 分	2.55 (0.64****)	2.9 (0.48)	2.64 (0.44)
	除去零分的平均分	2.95 (0.74)	3.13 (0.52)	2.85 (0.48)

**注：**\*本题得零分的学生对这部分知识基本处于前结构水平状态，从平均分上看，该部分学生人数不少，值得深思。

\*\*表示在本题中的小题号。

\*\*\*2015年10月和2016年4月的学考和选考都以2017届学生为主；2016年10月考试学考生以2018届为主，选考生以2017届为主；以后将类似于2016年10月考，并常态化。

\*\*\*\*表示该题在特定计算方式下的难度系数。

从3次题目的主干物质来看，均为二碳有机物，只有烷烃要求了解同系物，例如1604题（4）中的己烷；从知识结构来看，1510题主要为了解层次，1604题主要为理解层次，1610题则更侧重于运用层次，难度依次增大。

## （五）SEC一致性分析

国外对考试命题与课程标准的一致性分析的模式主要有三种：SEC、Webb、Achieve.

美国联邦政府和十几个地方州市都采用SEC模式进行考试命题与课程标准一致性的评价研究。SEC模式以“内容要素和认知层次”的二维矩阵构建为基础，波特（Porter）一致性系数 $P$ 表示，计算简单、应用广泛、适用性强。

$$P = 1 - \frac{\sum_{i=1}^n |X_i - Y_i|}{2}$$

其中： $n$ ——矩阵元总数目； $X_i$ —— $X$ 矩阵第 $i$ 个矩阵元； $Y_i$ —— $Y$ 矩阵第 $i$ 个矩阵元。 $P$ 的值在0~1范围区间， $P$ 值为0时，两研究对象间的一致性程度最小， $P$ 值为1时两研究对象间的一致性程度最大。

我们可以依据 $P$ 值的大小比较不同学考化学试题与《普通高中化学课程标准》间的一致性差异。学考化学试题命制与国家化学课程标准保持一致性又是极其重要的，不仅直接关系到学考化学试题命制的信度与效度，而且直接影响到课程预设和课程实施间的吻合度。基于以上原因，我们采用SEC定量一致性分析的方法，从“内容要素维度”“认知层次维度”“学科素养维度”三个方面来研究学考化学试题与课程标准之间的一致性，以期为新课程背景下的学考化学试题命制提供可资借鉴的实证依据。

## （六）国际测评项目经验

表3 国际教育测评经验

代表人物	成果
PISA（2015）	科学素养包括四个方面：科学知识、科学态度、科学能力和情境脉络。试题着重应用及情境化。受测学生必须灵活运用学科知识与认知技能，针对情境化的问题自行建构答案，因此能深入检视学生的基础素养。
TIMSS（2015）	科学测试的认知水平包括：了解、应用和推理。
NAEP（2009）	评价框架由四部分组成：学科领域、知与行能力（概念理解、科学研究、实际推理）、科学本质（科学史、思维习惯、探究方法和问题解决）、统一的概念（模型、系统、变化方式）。提出了4种认知类型：陈述性知识、程序性知识、图式知识和策略性知识。

以上成果对试题构造及学科能力表现的研究提供了国际视野。

## （七）国内相关教育考试文件

### 表4 国内有关教育考试评价改革文件

文件名称	成果
《国家中长期教育改革和发展规划纲要（2010—2020年）》（2010）	建立分类考试、综合评价、多元录取的考试招生模式。
国务院《关于深化考试招生制度改革的实施意见》（2014.9）	推行普通高校基于统一高考和高中学业水平考试成绩的综合评价多元录取机制。基于学业水平考试“3+3”模式形成。
教育部《关于全面深化课程改革，落实立德树人根本任务的意见》（2014.4）	将“发展学生核心素养体系”的研制与构建作为推进课程改革、深化发展的关键环节，落实“立德树人”目标。
教育部《关于普通高中学业水平考试的实施意见》（2014.12）	规定自主选考科目的学业水平考试计入高校招生录取总成绩；学业水平考试成绩合格，作为普通高中学业毕业及高中同等学力认定的主要依据。
教育部《中国学生发展核心素养》（2016.9）	学生发展的6大核心素养：人文底蕴、科学精神、学会学习、健康生活、责任担当、实践创新。
《高考考试大纲》（考试中心）（2017.12）	必备知识、关键能力、学科素养、核心价值的命题主旨。
《中国高考评价体系》和《中国高考评价体系说明》（2019.11）	确立了“一核”、“四层”和“四翼”的评价体系。

新高考改革确定将学业水平考试分为合格性考试（以必修课程要求为准）和学业自主选择计入高考录取成绩的等级性考试（以必修课程和选择性必修课程要求为准）。学业水平考试命题坚持以学科核心素养为导向，评价学生学科素养的发展状况和学业质量标准的达成程度。



# 教育考试题型建构的理论与方法

试题的类型是反映考试内容的形式，它服务于考试测量的目的和内容与要求，不同的题型，实现的功能不同。题型与考试的效度和信度有非常密切的关系，同时，大规模教育考试对学校的教学和教学训练方式有明显的“指挥棒”作用。学业水平考试应结合命题宗旨和目标，设计合适的题型，以实现对学生核心素养的测试。目前的化学高考试题有单选题和非选择题两种题型，其中非选择题的形式是填空题。这种碎片化的知识与能力考核形式极大地限制了学生的思维水平的发挥，也无法真实准确地考查学生的化学学科核心素养发展状况。

根据考生解题时思维方式的特点，高中学业水平考试题型可以分成7类：

**1) 再认型试题：**指对所学知识回忆、再认、辨认，即可回答试题。学业水平考试这类试题常加以“包装”，并跟STSE（Science, Technology Society and Enviroment）相联系。

**2) 推演型试题：**指对已学知识直接引用，再根据相似、类似、借用、模仿的规律举一反三，进行推理考察就可以回答的试题。

**3) 开发型试题：**指必须发掘已学知识的内涵，包括描述性和程序性知识，作较深层次开发，从中发现隐含的或自行总结归纳出某些规律和解题关键，据此再予以发散、推理才可以解决的试题。

**4) 重组型试题：**指将已经掌握的多种信息(知识块、信息块)予以筛选、提取、转换、倒置，特别是正向、逆向、横向重组后才能解决的试题。

**5) 结构反应型试题(也称信息迁移题)：**指必须将试题所给予的新信息吸收消化后重组成新的知识结构，然后再按再认型、推演型、开发型或重组型的思维方式才能解决的试题。

**6) 探究型试题：**指问题本身比较复杂，要细加分析，分别予以讨论才能得到完整答案的试题。或者是有多种可能解答还需要分别比较评价选择最佳答案的试题。

**7) 结构不良型试题（也称开放性试题）：**是指没有明确解决方法的问题。乔纳森把结构不良问题的解决过程总结为：理清问题及其情境限制，澄清、明确各种问题的角度、立场和利害关系，提出可能的解决方法，评价各种方法的有效性，对问题的表征和解法的监控，实施、监察解决方案，调整解决方案7个环节。由于此类试题结构上的不明确性，解法规则和答案上具有模糊性、开放性和多样性，且需要综合大量不同领域的知识解决问题，它有利于考核学生创新思维和解决问题能力。目前这种分类方法是在命题实践中根据能力测试的特点制定的，尚未定型。它作为调整全方位测试学科核心素养的试题组卷有一定的实践意义。

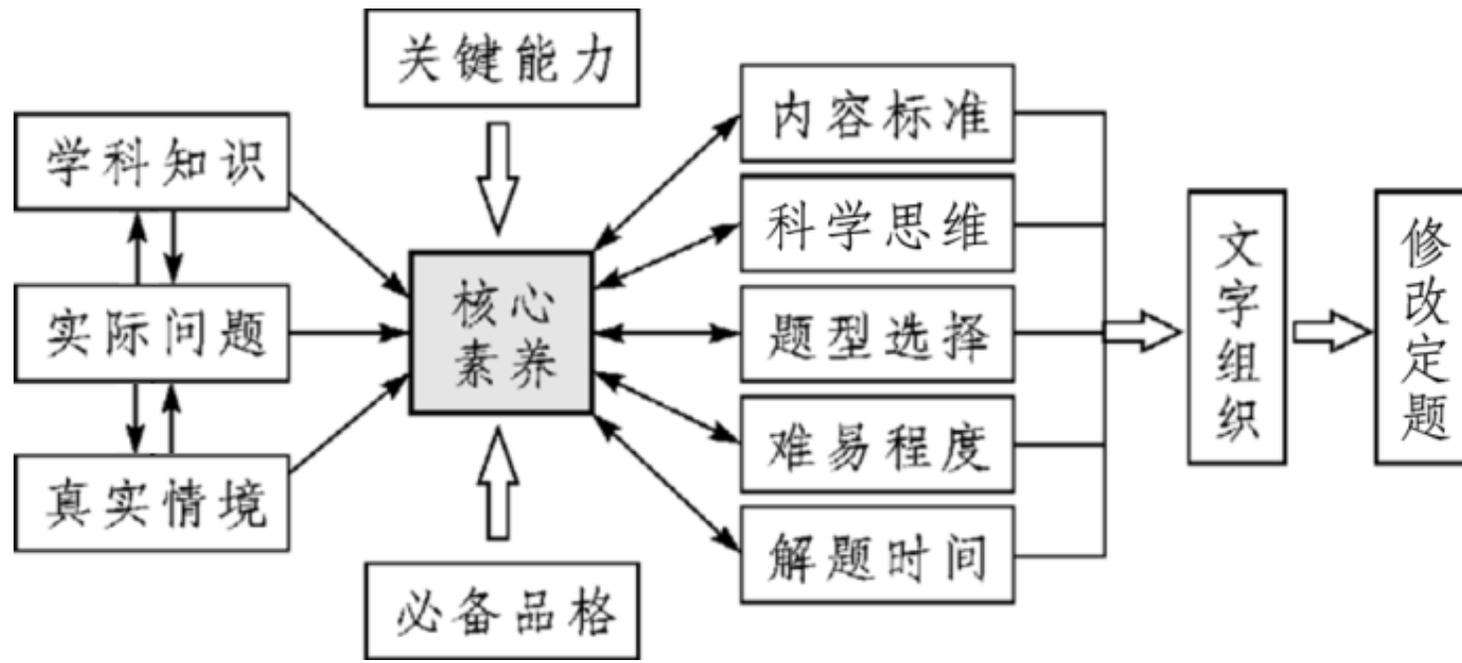


图2 其于学科核心素养为导向的命题框架

命题必须坚持以学科核心素养为导向，准确把握“素养”、“问题”、“情境”和“知识”4个要素在命题中的定位及相互关系。“情境”和“知识”同时服务于“问题”的提出和解决；“问题”“情境”“知识”三者之间存在密切的联系；情境的设计、知识的运用、问题的提出与解决应有利于实现对对学生核心素养的测试。同时要运用考试测量技术，呈现不同陌生度的问题情境，搭建丰富、生动的测试载体，从而确定试题。



构建中国特色教育考试评价体系

## （一）评价体系的创新

基于“素养为本”的考试命题要探索构建“一体、四层、四翼”的命题体系。

“一体”强调“立德树人、服务选才、导向教学”的核心立场。

## “四层”主要指：

**1) 必备知识：** 强调考查学生长期学习的知识储备中基础性、通用性知识，“学科本质”是最能反映学科核心素养的高质量的知识，它包括知识的产生和来源、事物的本质和规律、学科的思想与方法、知识的关系与结构、知识的作用与价值等意义系统。

**2) 关键能力：** 重点考查学生所学化学知识的运用能力，强调独立思考、分析问题和解决问题等学生适应未来不断变化发展社会的至关重要的能力。以化学为例，包括：化学语言和应用能力、接受和处理化学信息能力、化学方法和分析能力、化学实验及探究能力、化学计量及计算能力。

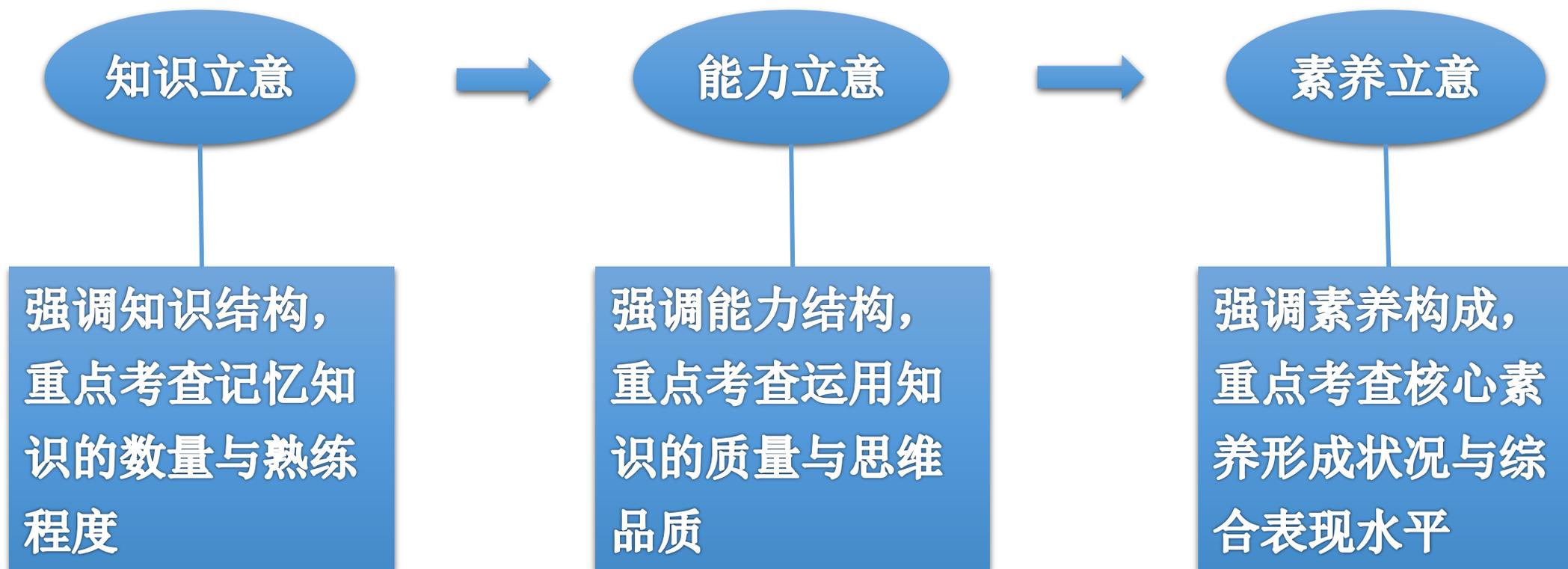
**3) 学科素养：** 要求学生能够在不同情境下综合利用所学知识和技能完成测试任务，具有扎实的学科观念和宽阔的学科视野，并体现出自身的实践能力、创新精神等内化的综合学科素养。

**4) 核心价值：** 要求学生能够在知识积累、能力提升和素质养成的过程中，逐步形成正确的核心价值观。

**“四翼”主要指基础性、综合性、应用性和创新性的要求。**

具体操作时，应以“核心价值”为引领，在“必备知识”基础上对“关键能力”和“学科素养”的全面覆盖。对不同思维方式、素养构成有别的考生进行全方位的考查。碎片化的知识与能力考核无法推动化学教育的发展，无法满足落实发展学生的核心素养这一根本目标。因此，应该设置相应的问题情境，增加实践型（如“化工流程图”）、探究型、信息迁移型试题比重，探索结构不良型试题、开放性试题等新题型命题方式，适当将填空式作答改为问答式、开放式作答，让学生充分地表达自己的思维状况和逻辑推理水平，同时可以检测学生的语言表达能力，为测查高层次的学科核心素养奠定基础。

## (二) 测量目标的转化



四

高考/学考试题难度的控制技术

试题难度设计与控制不仅是科学性问题，也是社会性问题。试题难度设计是指在题型设计时，编制出满足难度分布要求的试题。试题难度控制是指按照一定方法对试题的难易程度进行适当调适的过程。

试题“静态难度”是从试题的内部构成要素出发，利用认知任务分析的方法，对学生的认知水平、能力层级以及心理特征等进行检测的一种客观维度。（见表5）

试题“动态难度”则与考生群体能力强弱、试题新鲜度等有关，是一种相对难度。通常讲的高考试题（卷）的难度是考试后统计的结果，实际指的是“动态难度”。

**表5 理科试题“静态难度”的影响因素及量化标准（皇甫倩、王后雄，2017）**

难度因素	量化标准的说明
1.试题长度	阅读文本（含文字、符号、图表等）信息量的多少（单位：字符数）。
2.情境素材	试题所创设问题情境的新鲜度、繁简度、提示度、位置度等。
3.问题类型	题型对答案潜在的猜答的难度（封闭、半封闭、开放）。选择、填空、开放题。
4.知识点数	试题中所考查的知识点（含技能、方法）数目的多少（广度要求）。
5.能力层级	解题所涉及的能力层级水平（识记、理解、应用、评价）高低。
6.推理步骤	解题思维所涉及的推理步骤数（推理量）的多少。
7.运算步骤	计算题对数学技能的要求及运算的复杂程度（运算量）。
8.障碍点数	试题潜在的解题障碍点数（陷阱数、关卡数）及隐蔽条件的多少。
9.关联指数	以分步设问的各小题间是否循序渐进的、阶梯式的排列。

通过对影响难度因素的控制与难度实测数据进行比对分析，发掘试题结构与试题难度间的关联，对后续命题难度的调控具有一定的实践意义。

五

高考/学考双向细目表创新设计

双向细目表（two-way checklist）是一个测量的内容材料维度和行为技能所构成的表格，它能帮助命题者决定应该选择哪些方面的题目以及各类题目应占的比例。

考试命题双向细目表是一种考查目标（必备知识、关键能力、学科素养）和考试内容之间的关联表。双向细目表的制作应该同课程标准、考试说明的相关规定具有一致性。

命题的双向细目表要能一目了然地显示试卷的整体构想，同时，各道试题的考查内容、考查要求，以及在试卷中的位置（题号）也都记录在案。

随着时代的发展，我们可以对学考命题双向细目表进行改造，加入一些新元素，使双向细目表变得更加丰富，增强了对命题工作的指导性。以表6、表7是两种不同模式的学业水平考试双向细目表的框架。

表6 A学科学业水平考试命题双向细目表

题次	题型	分值	考点	能力要求				难度级别			预估难度
				A	B	C	D	容易	中档	稍难	
1											
2											
合计											

（“能力立意”双向细目表）

## 表7 B学科学业水平考试命题双向细目表

题号	章/节	考点	学科素养/分值					水平划分/分值				预估难度
			素养一	素养二	素养三	素养四	素养五	水平1	水平2	水平3	水平4	
1												
2												
合计												

（“素养立意”双向细目表）

双向细目表没有标准的固定的表格，实际运用时，往往可根据需要改变栏目的位置，变两个维度为多个维度的表格，展示试题（卷）整体的规划蓝图。

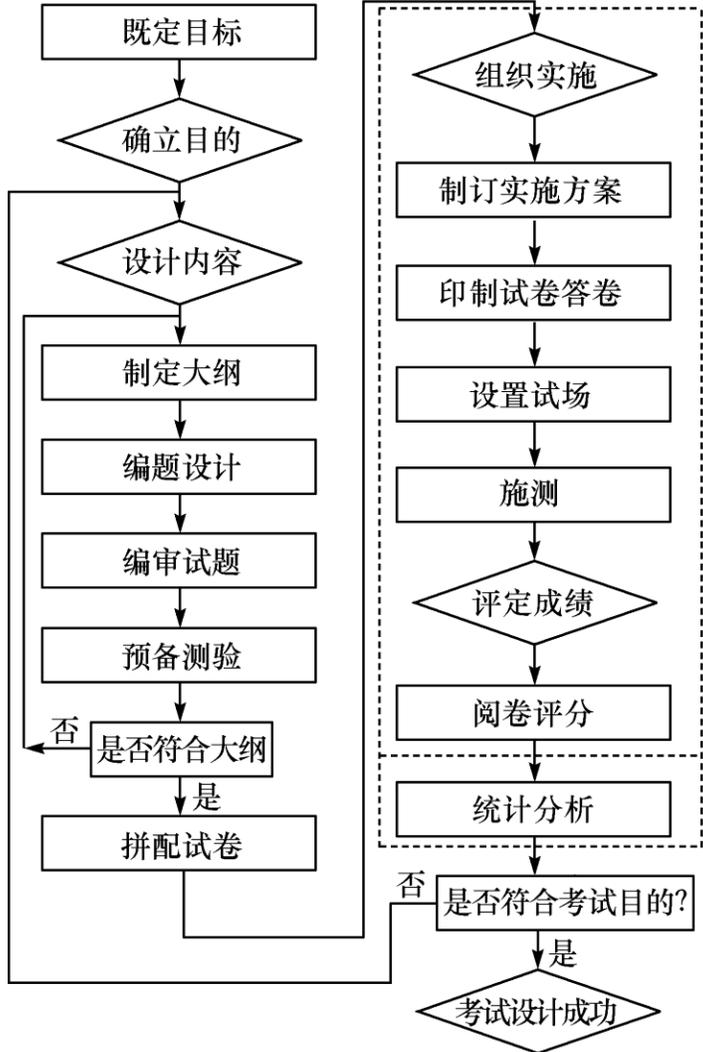
# 表8 高考理综化学命题双向细目表

题型	题号	知识组块	能力层次/分值				合计 分值
			识记	理解	应用	评价	
选择题(7-13)	7	STSE(化学史)	6				42分
	8	$N_A$ 的判断			6		
	9	化学常用计量			6		
	10	化学实验操作及检验			6		
	11	电化学原理及应用				6	
	12	原子结构与元素周期律			6		
	13	元素化合物				6	
实验综合题(26)	26(1)	仪器的识别及操作	3				14分
	26(2)	装置的连接、气密性检查			3		
	26(3)	氧化还原反应方程式书写				3	
	26(4)	实验设计与探究				3	
	26(5)	误差分析与探因				2	
无机综合题(27)	27(1)	核心化学方程式书写		2			14分
	27(2)	物质分离与提纯方法选择			2		
	27(3)	试剂及操作的作用		3			
	27(4)	离子的检验及洗涤干净的判断			3		
	27(5)	电子式的书写			2		
	27(6)	新情境下制备过程化学方程式				2	
反应原理题(28)	28(1)	还原性强弱判断			2		15分
	28(2)	$K_p$ 计算			3		
	28(3)	化学键与能量的关系			3		
	28(4)	平衡常数、速率计算			4		
	28(5)	溶液中离子浓度大小比较			3		
[选考题]物质结构与性质(37)	37(1)	基态原子电子排布式		2			15分
	37(2)	$\sigma$ 键、 $\pi$ 键数目确定		2			
	37(3)	杂化轨道类型、等电子体			3		
	37(4)	晶体的判断、解释事实			3		
	37(5)	晶胞化学式、配位数、密度计算			5		
[选考题]有机化学基础(38)	38(1)	分子式、结构式确定		3			15分
	38(2)	化学方程式		2			
	38(3)	反应类型、命名		3			
	38(4)	限定条件的同分异构体结构式的书写			2		
	38(5)	官能团的检验			2		
	38(6)	有机合成路线				3	
合计	12	40	9	44	52	10	115分

# 六

教育考试的命题流程及评价标准

# (一) 考试试卷设计的流程



教育考试试卷设计流程图

## （二）考试评价体系的设计

一级标目	二级标目	主要指标
A 考前设计	A1 考试目标	课程标准、考试范围、测量目的等。
	A2 命题蓝图	考试目标具体化——绘制全息式命题细目表。
	A3 试题编制	情景化设计，知识、能力、素养要求，难度合适，素材公平，原创、科学、规范，修改与定稿。
	A4 试题打磨	小范围试测，修改和定题。
	A5 试卷编制	试题组配、一致性分析、规范无误。
B 实质标准	B1 知识要素	与课标一致性，知识覆盖面广等。
	B2 能力要素	各种能力比例恰当，与题型匹配適切。
	B3 素养要素	各方面素养比例适当，符合学业质量标准要求。
	B4 难度要素	试卷难度符合测试目标及群体的要求。
	B5 题型要素	各类题型的比例恰当，具有良好的导向作用。
	B6 长度要素	试卷的题量、文字及字符数量与考试时间匹配。
	B7 公平要素	选材不产生城乡、性别、民族、地域等偏见。
	B8 其他要素	全卷分数合理、考试时间恰当，考虑试卷等值。

## （二）考试评价体系的设计

一级标目	二级标目	主要指标
C 考试实施	C1 施测准备	考试宣传、考场编排、试卷印制及保管等。
	C2 现场施测	考场管理有序，考风、考纪严明，监考职责明确。
	C3 阅卷评分	集中封闭、安全保密、评分细则具有可操作性。
	C4 分数合成	登分、核分、统分工作程序规范、流畅。
	C5 诊断评价	科学使用考试统计结果，诊断、提升教学质量。
D 考后评价	D1 总体描述	考情总体分析，教学质量诊断，基于“证据”教学质量评价。
	D2 考试难度	考试难度系数，各题实际难度与预测难度差异分析。
	D3 试题区分度	各类试题在不同群体（好、中、差）反应特征。
	D4 考试信度	测量的一致性程度指标，一般要求0.90以上。
	D5 考试效度	测验对它所要测量的特质准确测量的程度。
	D6 其他分析	考试导向性、科学性、公平性分析，试卷差错率等。

### （三）命题审核的一般要求

#### ×× ××考试××学科命题审核表

科目代码：\_\_\_\_\_

科目名称：\_\_\_\_\_

类别	序号	审核项目	命题人自检 (签名确认)	命题组长复核 (签名确认)
学习文件	1	已经学习并了解学校有关命题的质量和纪律要求		
	2	已经签署《××××自命题科目命题工作责任书》，已经做好命题的一切准备工作。		
基本信息	3	已命制完成试题，现在开始检查：试卷标题年份是否为“二〇一八”、试题类型是否标准明确、“考试科目及名称”是否与招生简章公布的一致。		
	4	卷面各题分值标注及总分值是否规范、准确（总分值是否符合试题满分要求，具体分值见科目设置表）		
试题编排	5	是否采用正确的纸型规格		
	6	题序、题号是否正确、规范		
	7	字体、字号设置是否符合规范（各题标题字号为宋体加粗小4号字，题干字号为标准宋体5号字）		
	8	页码标注是否正确（应设置为“第×页 共×页”）		

试卷 科学	9	卷面语言是否通顺、符合语法，无歧义，无错别字		
	10	题目是否可解，无科学性、政治性错误		
	11	该科目是否按要求命制了一套或两套试卷（AB卷）		
	12	该科目两套试卷（AB卷）的题量、难度和试题水平是否接近或相当，原则上题目不得重复。		
试题 印制 装袋	13	清点从院、校领取的该科目试题信封（AB卷两套信封），密封笺是否完好		
	14	核对信封上书写的科目代码和名称是否与命题科目和名称完全一致		
	15	将试卷对折后小心放入试题专用信封中，注意保持试卷整洁，不要污损试题		
	16	检查该科目两套试卷（AB卷）是否分别放入两个AB试题信封		
	17	将试题袋封口，贴密封签		
	18	命题组长在试题封口处签名。		
	19	将试题信封和责任书交试题接收人员保存。		
	20	所有交接人员签名确认		

分管领导签字（盖章）：\_\_\_\_\_年\_\_\_\_月\_\_\_\_日

|| [ 七 ] ||

高考/学考评分标准规则的制定

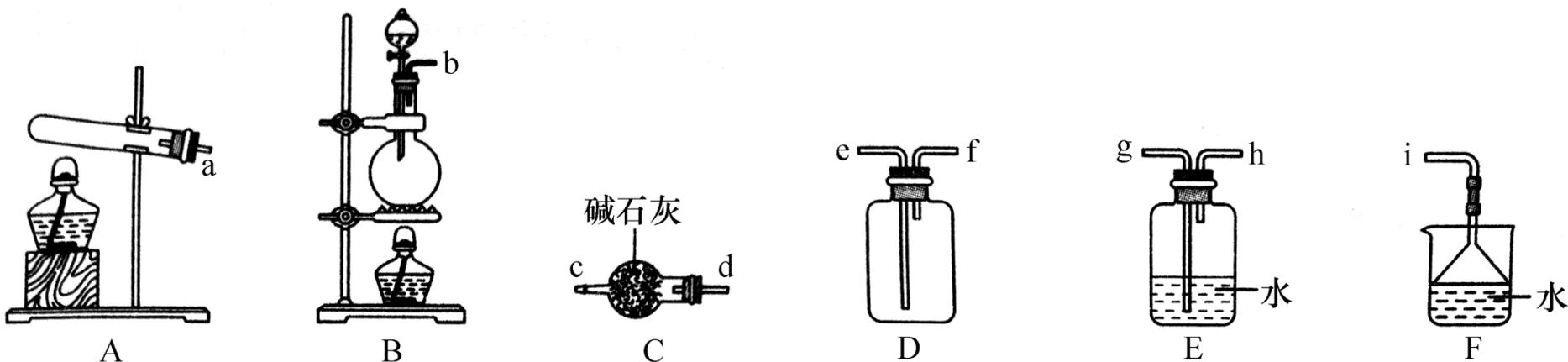
评分标准是为了方便阅卷教师分数评定时相对客观的准则，控制阅卷教师评分误差的评判细则。评分标准的制定越全面、越细致，评分误差就会越少。

- (1) 认真编写每道试题的答案及详实的评分标注。
- (2) 合理确定采分点，逐段赋分，给出评分标准。
- (3) 尽量平衡不同解法的评分标准，严控评分误差。

[示例] (2016年全国卷 I)

26. (14分) 氮的氧化物 ( $\text{NO}_x$ ) 是大气污染物之一, 工业上在一定温度和催化剂条件下用  $\text{NH}_3$  将  $\text{NO}_x$  还原生成  $\text{N}_2$ 。某同学在实验室中对  $\text{NH}_3$  与  $\text{NO}_2$  反应进行了探究。回答下列问题:

(1) 氨气的制备



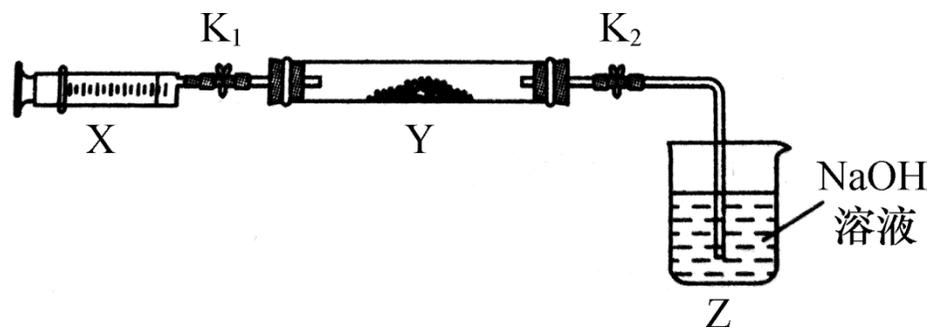
①氨气的发生装置可以选择上图中的\_\_\_\_\_，反应的化学方程式为

\_\_\_\_\_。

②欲收集一瓶干燥的氨气, 选择上图中的装置, 其连接顺序为: 发生装置  
→\_\_\_\_\_ (按气流方向, 用小写字母表示)。

## (2) 氨气与二氧化氮的反应

将上述收集到的 $\text{NH}_3$ 充入注射器X中，硬质玻璃管Y中加入少量催化剂，充入 $\text{NO}_2$ （两端用夹子 $\text{K}_1$ 、 $\text{K}_2$ 夹好）。在一定温度下按下图装置进行实验。



操作步骤	实验现象	解释原因
打开 $\text{K}_1$ ，推动注射器活塞，使X中的气体缓慢通入Y管中	①Y管中_____	②反应的化学方程式_____
将注射器活塞退回原处并固定，待装置恢复到室温	Y管中有少量水珠	生成的气态水凝聚
打开 $\text{K}_2$	③_____	④_____

## [答案及分值]

(1) ①A  $2\text{NH}_4\text{Cl} + \text{Ca}(\text{OH})_2 \xrightarrow{\quad} 2\text{NH}_3\uparrow + \text{CaCl}_2 + 2\text{H}_2\text{O}$  (或B  $\text{NH}_3 \cdot \text{H}_2\text{O} \xrightarrow{\quad} \text{NH}_3\uparrow + \text{H}_2\text{O}$ )  
(1分, 2分, 共3分)

②dcfei (3分)

(2) ①红棕色气体慢慢变浅 (2分)

② $8\text{NH}_3 + 6\text{NO}_2 \xrightarrow{\quad} 7\text{N}_2 + 12\text{H}_2\text{O}$  (2分)

③Z中NaOH溶液产生倒吸现象 (2分)

④反应后气体分子数减少, Y管中压强小于外压 (2分)

## [评分细则]

(1) ①书写化学方程式时, 没写“ $\xrightarrow{\quad}$ ”“ $\uparrow$ ”等反应条件及气体符号的扣1分, 未配平方程式的扣1分。若方程式出现错误的化学式时, 不得分。

②按参考答案给分, 与参考答案不相符的不得分。

(2) ①描述现象时若出现“红棕色气体消失”不扣分。

②化学方程式中未写“催化剂”的扣1分, 未配平方程式的扣1分。

③描述现象时若出现“烧杯中的溶液沿玻璃管进入硬质玻璃管Y中”不扣分, 若只写“倒吸”的扣1分。

④若只答“Y管中压强小于外压”的扣1分, 若答“反应后气体物质的量减小, Y管中的气压小于大气压”不扣分。



# 高中学考试卷评价指标体系设计

学业水平考试试卷评价量表的建立和使用按照以下思路进行：  
一是试卷整体评价与个性分析相结合，对学业水平考试试卷具备的功能进行全面评价，同时关注不同学科试卷的特点；二是第三方评价（非命题者和非被试对象的考生），站在第三方的角度开展评价，客观分析试卷是否符合学业水平考试的要求；三是定量评价（客观数据）与定性评价相结合，评价教师运用自己教学积累的经验分析统计结合，评价教师运用自己教学积累的经验分析统计结果，探究影响教学的种种因素和产生问题的原因，提出改进建议。（朱宁，2008年）

## 表9 高中学业水平考试评价测量指标体系

项目		内容
①学科整体分析 ②题目逐题分析 ③细分小题分析	指标 汇总	平均分，最高分，最低分，难度，区分度，标准差，差异系数，鉴别指数，信度系数，峰度，偏度。
	图类 汇总	总分分布曲线图，分数分布曲线图，难度曲线图，分组难度曲线图。
	表类 汇总	试卷总分次数分布表，等级成绩分析，题目或题组等级得分率分析表，客观题选项等级分析表，客观题选项分组分析表，主观性试题得分分组分析表，G1~G5分组分析或G1~G7分组分析*。
学科相关分析		主客观题组块相关分析表，知识组块相关分析表，能力组块相关分析表，基础性试题和发展性试题相关分析表，教材板块相关分析表。

\*注：G1~G5分组或G1~G7分组是按照考生成绩由低到高排列，按照总人数平均分成5组或7组。同时还应设计高中学业水平考试试卷命题定性评价表，规范试题质量与分平标准。（任涛等，2017年）

## 表10 高中学业水平考试试卷命题定性评价表

评价维度	内容标准
① 试卷结构技术指标科学、合理	考查题型比例配置适当，各种难度的试卷比例适当，试卷长度合适等。
② 试题符合相应题型的编制要求	试题测试目标明确，立意、情境、设问科学、试题情境素材选择符合公平性准则，试题表述方式合理、有效等。
③ 参考答案和评分标准科学、准确	参考答案科学、准确、没有争议，评分参考的赋分合理，便于操作，评分参考有利于控制评分误差等。
④ 试卷考查的知识内容满足要求	考查内容具有合理的覆盖面和比例，教材体系中不同知识板块考查比例合适，注重教材差异，突出主干、核心观念的考查，素材背景的导向良好等。
⑤ 试卷考查的能力要求全面合理	不同试题考查的能力分类分层均有体现，试卷注重考查基础或发展性能力，联系实际问题能力、科内综合或跨科综合能力突出考查等。
⑥ 试卷考查的学科素养全面合理	全面考查学生学科核心素养的发展状况和学业质量标准的达成程度。情境的设计、知识的运用、问题的提出与解决均应有利于实现对学生核心素养的测试。
⑦ 试卷测量指标（难度、区分度等）符合要求	符合既定的设计目标，内容（知识）难度，能力难度，题目位置难度，较好的区分度等。
⑧ 试卷符合科学规范的要求	格式规范，学科语言和符号的规范等。

结合具体学科可以设计试卷的评价角度和具体指标。（见表11）

**表11 高中化学学业水平合格考试试卷评价表（示例）**

项目	内容	题号	分值（分）
题型分类	客观性试题（选择题）	1-25	50
	主观性试题（非选择题）	26-32	50
知识组块	化学科学与实验探究	4、5、18、23、29	18
	常见的无机物及其应用	2、6、7、20、30	18
	物质结构基础	9-11、21、24、27	14
	化学反应规律	13-15、22、28	12
	简单的有机化合物及其应用	3、16、17、31	16
	化学与社会发展	1、8、19、26	10
	化学计算	12、25、32	12
能力组块	学习理解	1-20、24、26、27	46
	应用实践	21-23、25、28、30-32	44
	迁移创新	29	10
学科素养	宏观辨识与微观探析	9-11、16、17、24、27、30	30
	变化观念与平衡思想	7、13-15、20、21、26、31	18
	证据推理与模型认知	2、3、6、12、22、25、28、32	22
	科学探究与创新意识	4、5、18、23、29	18
	科学态度与社会责任	1、8、9	6
题型分析	选择题	1-25	50
	填空与简答题	26-28	12
	化学实验与探究题	29	10
	无机综合题	30	10
	有机综合题	31	10
	综合计算题	32	8

此外，在考试结束后可用SPSS软件或EXCEL进行全样本的考生数据分析，获得学业水平考试试卷的数据报告（见表12），可以进行基于标准的定量数据分析。

**表12** ××××年高中学考C学科数据分析表（示例）

统计类别	题目或题组	人数	满分	最高分	最低分	平均值	标准差	差异系数	得分率	区分度
总体	C学科全卷	45192	100	100	3	74.37	12.02	0.16	0.74	——
题型分类	客观题	45192	75	75	2	61.73	8.90	0.14	0.82	0.96
	主观题	45192	25	25	1	12.63	4.25	0.34	0.51	0.82
知识组块	知识块1	45192	3	3	0	2.98	0.22	0.08	0.99	0.17
	知识块2		6	6	0	5.38	1.27	0.24	0.90	0.55
能力组块	能力1	45192	11	11	0	10.01	1.82	0.18	0.91	0.64
	能力2		9	9	0	6.71	1.49	0.22	0.75	0.28
学科素养	素养1	45192	24	24	0	7.68	4.14	0.54	0.32	0.73
	素养2		28	28	0	20.58	3.96	0.19	0.73	0.82
教材分布	必修1	45192	26	26	0	15.32	3.81	0.25	0.59	0.76
	必修2		25	25	0	20.03	3.70	0.18	0.80	0.82

考试后按照表12所列的项目撰写学科评价报告，一方面可使评价的过程标准量化、有据可依，增加了评价报告的规范性和可读性，也能使教育管理者或命题教师一目了然地熟悉试卷评价诊断反映出的问题，便于后续命题改进，同时也是对教师教学质量和学生学业质量进行诊断评估的重要依据，是提升我国学业质量的重要保障。

# 九

“素养为本” 试题分析模式示例

**[例2]**（2018年高考理科综合全国 I 卷第27题）焦亚硫酸钠（ $\text{Na}_2\text{S}_2\text{O}_5$ ）在医药、橡胶、印染、食品等方面应用广泛。回答下列问题：

（1）生产 $\text{Na}_2\text{S}_2\text{O}_5$ ，通常是由 $\text{NaHSO}_3$ 过饱和溶液经结晶脱水制得。写出该过程的化学方程式\_\_\_\_\_。

（2）利用烟道气中的 $\text{SO}_2$ 生产 $\text{Na}_2\text{S}_2\text{O}_5$ 的工艺如图2所示：

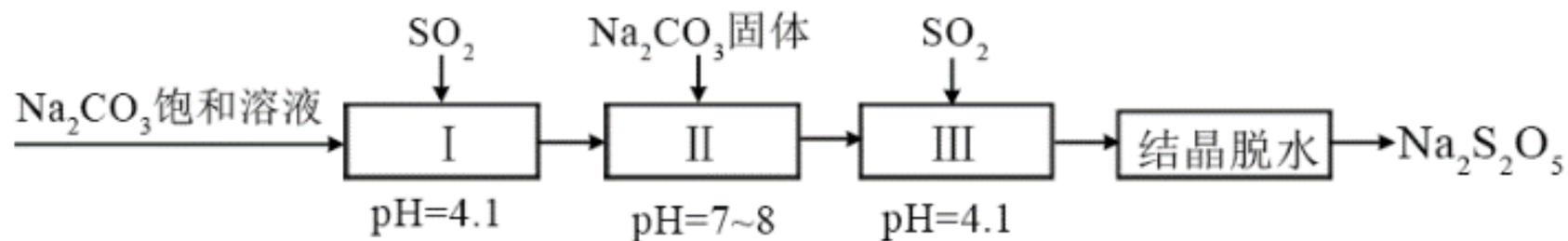


图2

- ①  $\text{pH}=4.1$ 时，I 中为\_\_\_\_\_溶液（写化学式）。
- ② 工艺中加入 $\text{Na}_2\text{CO}_3$ 固体、并再次充入 $\text{SO}_2$ 的目的是\_\_\_\_\_。

(3) 制备 $\text{Na}_2\text{S}_2\text{O}_5$ 也可采用三室膜电解技术，装置如图3所示，其中 $\text{SO}_2$ 碱吸收液中含有 $\text{NaHSO}_3$ 和 $\text{Na}_2\text{SO}_3$ 。阳极的电极反应式为\_\_\_\_\_。电解后，\_\_\_\_\_室的 $\text{NaHSO}_3$ 浓度增加。将该室溶液进行结晶脱水，可得到 $\text{Na}_2\text{S}_2\text{O}_5$ 。

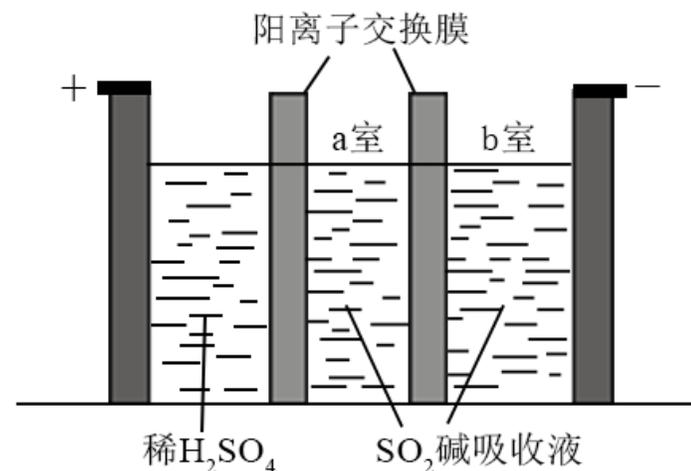


图3

(4)  $\text{Na}_2\text{S}_2\text{O}_5$ 可用作食品的抗氧化剂。在测定某葡萄酒中 $\text{Na}_2\text{S}_2\text{O}_5$ 残留量时，取50.00 mL葡萄酒样品，用 $0.01000 \text{ mol}\cdot\text{L}^{-1}$ 的碘标准液滴定至终点，消耗10.00 mL。滴定反应的离子方程式为\_\_\_\_\_，该样品中 $\text{Na}_2\text{S}_2\text{O}_5$ 的残留量为\_\_\_\_\_ $\text{g}\cdot\text{L}^{-1}$ （以 $\text{SO}_2$ 计）。

本题是以焦硫酸钠的生产为背景的化学工艺流程题，在问题设置时围绕“预处理”、“核心反应”、“提纯分离”展开，主要考查化学方程式的书写、具体工艺措施的解释，并结合电化学制备方法及氧化还原滴定考查电极反应方程式的书写、离子反应方程式书写及残留量计算。这些内容主要考查了学生运用化学原理解决实际问题的能力，其主要涉及的化学学科核心素养为“证据推理与模型认知”，表13所示为本题各小题考查的核心素养的维度及级别水平。

表13 例2各小题核心素养维度及素养级别

题号	考查的核心素养	核心素养的表现	素养级别
(1)	宏观辨识与微观探析	(由题干反应物和生成物信息书写方程式)能运用化学符号描述常见简单物质及其变化	水平1
(2) ①	证据推理与模型认知	(由信息推理 I 的化学成分)能从物质及其变化的事实中提取证据,对有关的化学问题提出假设,能依据证据证明或证伪假设	水平1
(2) ②	证据推理与模型认知	(根据流程图中相关信息推理通入二氧化硫的目的)能从宏观和微观结合上收集证据,能依据证据从不同视角分析问题,推出合理的结论	水平2
(3)	证据推理与模型认知	(由三室膜电解技术示意图推测电极反应式及电解室浓度变化)能从宏观和微观结合上收集证据,能依据证据从不同视角分析问题,推出合理的结论	水平2
(4)	科学探究与创新思维	(由实验过程及数据,书写滴定反应方程式,计算焦硫酸钠残留量)收集实验证据,基于现象和数据进行分析并得出结论	水平3



# 开展教育考试与评价的实践研究

# 结题证书

课题类别：全国教育考试“十一五”科研规划课题（2006JKS10）

课题名称：新课程下高考质量评价标准体系及其实施研究

课题负责人：王后雄

## 教育部人文社会科学 结项证

项目类别：规划基金项目

项目名称：国家考试制度公平标准及评价体系研究

负责人：王后雄

主要参加人：

批准号：09YJA880049

本项目经审核准予结项，特发此证。

证书编号：2012JXZ0083

教育部社会科学司

2012年02月16日

## 教育部人文社会科学研究项目

# 结项证书

项目类别：规划基金项目

项目名称：国家教育考试公平监测指标体系及保障机制研究

负责人：王后雄

主要参加人：孙建明 余澄 詹先君 王世存 李木洲  
朱晨辉

批准号：14YJA880066

本项目经审核准予结项，特发此证。

证书编号：2017JXZ1143

教育部社会科学司

2017年07月13日



王后雄，湖北省特级教师，享受国务院政府特殊津贴的专家，从事大学教学与考试研究工作。现任华中师范大学考试研究院院长、教师教育学院副院长、化学教育研究所所长。其首创的“化学教学诊断学”“中学化学目标控制教学法”在全国1000多所中学实验推广，效果显著。先后被授予全国劳动模范、全国改革“十佳”教师，多次受到党和国家领导人的接见。为10多个省市大型考试的主要命题人，是湖北省新高考方案制订专家组核心成员之一。

## 新高考、新趋势、新思路、新策略

——访华中师范大学考试研究院院长王后雄教授

◎本刊记者□陈兰枝 夏豪杰

**【编者按】**自新中国成立以来，我国高考历经了数次变革，从高校单独招考到“大文大理”设科，从高中毕业会考与高校招生考试两考并置再到“3+X”设考，但高考的“指挥棒”效应一直没有从根本上得到改变。2014年启动的新高考改革是历史上“力度最大”的一次高考变革，承载了社会、家长和学生的多重希望。提起王后雄，在大多数人眼中他是被誉为与任志鸿、薛金星齐名的教辅大师。实际上，他还有一个身份鲜为人知，那就是作为华中师范大学考试研究院院长，他是师承我国考试学科创始人廖平胜教授的考试学专家。近年他主持了多项与高考相关的国家级课题研究，为了掌握本轮高考改革的实际情况，他带领研究团队多次到新高考试点地区，深入高中学校走访调研，取得了许多阶段性研究成果。新高考如何考，高考分数如何算，学生选科如何选，班级管理如何管，学生、家长、高中学校如何应对，是目前大家都非常关心的问题，带着这些问题，本刊记者对王后雄教授进行了访谈。

**记者：**王教授，您好！2014年9月，国务院正式印发《关于深化考试招生制度改革的实施意见》（以下简称《意见》），拉开了我国新一轮高考改革的序幕。根据《意见》的部署，上海和浙江率先启动高考综合改革试点，目前两地学生已经进入高三年级，全社会极其关注。从历次高考改革的发展历程来看，您认为此次新高考改革的主要目的或动因是什么？

**王后雄：**高考改革一直都是我非常关心的教育课题。我曾经申请了三个国家级课题：全国教育考试“十一五”科研重点课题——新课程下高考质量

评价标准体系及其实施研究、教育部人文社科基金项目——国家考试制度公平标准及评价体系研究和教育部人文社科基金项目——国家教育考试公平监测指标体系及保障机制研究。在这三个课题的研究中，我对中国的高考发展历程有了更加深刻的认识。新中国成立以来，高考改革大致经历了五个阶段，每个阶段都有不同的考试模式。

第一个阶段的时间大致是1949—1952年。当时，新中国刚刚成立，许多制度还不完善，很多高等院校是从国民政府接收过来的，考试还沿袭以前的做法，即基本采取高校单独招考模式。这种模式

2004—2013年《中国考试》以第一作者发文最多的是18篇，依据普赖斯公式  $M=0.749\sqrt{N_{max}}$ （其中  $M$  为高频阈值， $N_{max}$  表示论文被引频次最高值）<sup>[78]</sup>，计算得到  $M=3.17$ ，选择发表4篇或4篇以上论文的作者为《中国考试》的核心作者候选人。

2004—2013年《中国考试》作者发文被引频次累积最高为99次，依据普赖斯公式确定核心作者候选人的发文累积最低被引量为： $M=0.749\sqrt{N_{max}}=0.749\sqrt{99}=7.45$ ，即发文累积最低被引8次或8次以上的作者为《中国考试》的核心作者候选人。

综合上述两个因素：发文4篇或4篇以上、发文

累积被引8次或8次以上的作者即为《中国考试》10年来的核心作者。经统计，核心作者有28位（见表2），共计发文184篇，占2004—2013年《中国考试》刊发研究论文总量的13.8%；核心作者发文被引频次共计809次，占总被引频次的30.5%。核心作者的发文数与文献计量学中核心作者发文数应占发文总量20%的下限有一定差距。<sup>[9]</sup>目前看来，《中国考试》尚未形成强大稳定的核心作者群，有待于进一步发展。

核心作者中，归属考试机构的作者有21位（其中教育部考试中心的作者有11位），高等院校的作

表2 《中国考试》2004—2013年刊文的核心作者

序号	作者	所在单位	以第一作者发文篇数(篇)	累积被引频次(次)
1	戴家干	教育部考试中心	10	99
2	王后雄	华中师范大学	17	61
3	张晋军	国家汉办	10	59
4	马世晔	教育部考试中心	4	50
5	柳博	教育部考试中心	6	48
6	辛涛	北京师范大学	4	47
7	陈睿	教育部考试中心	5	45
8	王蕾	教育部考试中心	13	39
9	谢小庆	北京语言大学	10	36
10	关丹丹	教育部考试中心	10	32
11	王志武	湖北省教育考试院	10	29
12	胡向东	湖北省教育考试院	5	28
13	赵守盈	贵州师范大学	5	25
14	王和军	教育部考试中心	4	23
15	李化德	西南政法大学	7	21
16	余扬	陕西省考试管理中心	6	21
17	张亚群	厦门大学	5	17
18	赵海燕	北京教育考试院	6	15
19	韩宁	教育部考试中心	5	14
20	刘远我	人事部人事考试中心	5	14
21	李光明	教育部考试中心	5	13
22	李金波	浙江省教育考试院	5	13
23	姜钢	教育部考试中心	4	12
24	王晓华	上海市教育考试院	6	12
25	刘炳贵	江苏省教育考试院	4	11
26	黄继晏	江西省教育考试院	4	9
27	任子朝	教育部考试中心	5	8
28	张敏强	华南师范大学	4	8
合计			184	809

1

HOUXIONG WANG

## Research on the Influence of College Entrance Examination Policies on the Fairness of Higher Education Admissions Opportunities in China

*Abstract: College entrance examination policies are the criteria developed by the government, institutions of higher education, and community organizations for the authoritative distribution and regulation of higher education opportunities. College entrance examination policies are the most immediate and most important factor affecting equality of access to higher education. In the thirty years since the restoration of the college entrance examination, policy discussions and empirical studies exploring fairness issues*

---

English translation © 2011 M.E. Sharpe, from the Chinese text, “Zhongguo gaokao zhence dui gaodeng jiaoyu ruxue jihui gongpingxing yingxiang yanjiu.” Translated by Laura and David Truncellito.

Houxiong Wang is a dean at the Academy of Assessment of Central Normal University, Wuhan, as well as a professor and advisor to Ph.D. students. He conducts research on education examinations and evaluations as well as curricula and pedagogy.

This study was subsidized by the Ministry of Education Humanities and Social Sciences research topic “National College Testing System Fairness Indicator and Evaluation System Research” (topic approval no. 09YJA880049) stage results.

---

RESEARCH ARTICLE

WANG Houxiong

## Access to Higher Education in China: Differences in Opportunity

© Higher Education Press and Springer-Verlag 2011

**Abstract** Access to higher education in China has opened up significantly in the move towards a mass higher education system. However, aggregate growth does not necessarily imply fair or reasonable distribution of opportunity. In fact, the expansion of higher education has a rather more complex influence on opportunity when admissions statistics are viewed by geographical region, rural and urban environment, social class, type of school, gender, and ethnicity. Since 1999, gaps in access opportunities have generally diminished, especially in terms of the urban-rural dimension. Efforts to increase university admission rates for women and minorities have resulted in significant progress. However, the gap in university admission between different social classes has been closing more slowly. Children from more advantaged backgrounds have more chances to study at key universities, and differences in access between provinces are still considerable. Social class polarization in secondary school is still a serious issue. Such problems at high school level directly lead to the accumulation and continuation of a gap in opportunities to access higher education. While agreeing that the aggregate growth of higher education provision in China is a positive development, we also strive to improve equality of opportunity.

**Keywords** mass higher education, access to higher education, differences in opportunity

---

## Introduction

After 30 years of reform of the college entrance examination, China now has a college entry system with the unified national college entrance examination as

---

Received April 14, 2010

WANG Houxiong (✉)

The Academy of Assessment, Huazhong Normal University, Wuhan 430079, China,  
E-mail: 67127513@163.com

1

HOUXIONG WANG

## A Commentary on China's New Curriculum and the Programs to Design Subjects for the College Entrance Examination

**Abstract:** *Designing and reforming the subjects on the college entrance examination, based on the new curriculum, are the focal point and also the most difficult aspect of entrance exam reform. The entrance exam subjects programs instituted in over ten "subject reform" regions in China, including the provinces of Shandong, Ningxia, Guangdong, Hainan, and Jiangsu, each possess unique characteristics; each has also given rise to a number of issues that urgently require resolution. In examining the present trend toward*

English translation © 2013 M.E. Sharpe, from the Chinese text, "Zhongguo gaokao zhence dui gaodeng jiaoyu ruxue jihui gongpingxing yingxiang yanjiu." Translated by Laura and David Truncellito.

Houxiong Wang is a dean at the Academy of Assessment of Central Normal University, Wuhan, as well as a professor and advisor to Ph.D. students. He conducts research on education examinations and evaluations as well as curricula and pedagogy.

This study was subsidized by the Ministry of Education Humanities and Social Sciences research topic "National College Testing System Fairness Indicator and Evaluation System Research" (topic approval no. 09YJA880049) study results.



Cite this: DOI: 10.1039/c6rp00076b

## Research on evaluation of Chinese students' competence in written scientific argumentation in the context of chemistry

Yang Deng\* and Houxiong Wang\*

Attending to practice has become a significant topic in science education today. As scientific argumentation is a typical form of scientific practice as well as an important educational practice, more and more attention has been paid to it by science education researchers. Evaluating students' competence in scientific argumentation is one of the most important research topics, but in China, science researchers seldom concentrate on it because the diverse educational values of scientific argumentation need to be further understood. The present study sought to examine the performance of Chinese students participating in written scientific argumentation in the context of chemistry. After clarifying the conception of scientific argumentation in science education, and comparing the evaluation criteria in domestic and international science education research, written scientific argumentation tasks in the context of chemistry were designed and criteria for their evaluation were constructed and improved. In total five tasks were designed for evaluation. All of the five tasks were aimed at evaluating students' competence of selecting (or putting forward) claims, evidence and warrants. In addition, two tasks investigated the competence of refuting arguments. The general criteria for evaluation was constructed according to the four dimensions of scientific argumentation, they were the structure components, the content quality, the logic of justification and language. For each task, content criteria and performance criteria for evaluation were constructed. After analysis and improvement of the criteria based on two pilot tests and the Rasch model, it was obvious that the criteria met the standards, effectively and credibly, for this study on the assessment of students' competence in written scientific argumentation. The number of students who participated in the formal test was 578 (304 males and 274 females). Through this kind of evaluation, this study found that the students' competence in written scientific argumentation was generally weak, and was influenced by some factors. Specifically, firstly, the students could put forward claims and evidence more easily than warrants and rebuttals. Secondly, the specific tasks had an influence on the performance of the students in written scientific argumentation. In regard to other factors, gender did not influence the students' competence in written scientific argumentation, but the grade level and school level were key factors. The students' competence in written scientific argumentation at grade level four and three other school grade levels were significantly different. Finally, some changes to the Chinese chemistry curriculum were proposed based on the results of this study.

Received 29th March 2016,  
Accepted 1st November 2016

DOI: 10.1039/c6rp00076b

www.rsc.org/cerp

## Introduction

Attending to practice has become a significant topic in science education today. As scientific argumentation is a typical form of scientific practice as well as an important educational practice (Driver *et al.*, 2000; Erduran *et al.*, 2004; Erduran, 2007), and several new instructional methods and curricula have been developed over the last decade to help students to acquire the competence needed to participate in scientific argumentation

(Zemba-Saul, 2009; Berland and Reiser, 2011; Walker *et al.*, 2011; Cavagnetto and Hand, 2012; Foong and Daniel, 2012; Hong *et al.*, 2013), more and more attention has been paid by science education researchers on evaluating students' competence in scientific argumentation (Kelly and Takao, 2002; Sandoval and Millwood, 2005; Ryu and Sandoval, 2012; Sampson *et al.*, 2012; Wu and Tsai, 2012; Mendonca and Justi, 2014). These types of studies can help us to examine how these new instructional methods and curricula work. Written argumentation is a form of argumentation which needs written language, it plays an important role in science because science can be treated as a form of language expression, especially the

Department of Chemistry, Central China Normal University, Wuhan, Hubei 430079, China. E-mail: yangdeng@mail.ccnu.edu.cn, 67127513@163.com

## 高考命题公正的现实困境与两难选择<sup>\*</sup>

王后雄

**[摘要]** 公正是高考命题各个环节必须考虑的现实问题。人们在追求命题公正实践中又面临公正的两面性、相对性和复杂性。高考试卷命制中,知识与能力的关系、试题内容、试卷难度、题型设计、试卷组配、试题情景等问题都关涉公正的两难选择。实现高考命题公正的基本途径是牢固树立考试公正理念;建立命题公正性审查制度;加强国家考试题库的建设;着力提高命题者的公正素养。

**[关键词]** 高考命题;公正性;科学性

**[作者简介]** 王后雄,华中师范大学考试研究院考试中心主任、教授 (武汉430079)

高考是我国高校科学选拔人才的大规模选拔性考试,因此,高考对公正公平性原则进行操作的第一要务,就是确定科学合理的选拔标准。它在一定程度上摒弃了权力、出身和人际关系等因素对教育公正的干扰,体现了客观性和公正性,保证了个人凭才能平等地接受高等教育的权利。然而,在“分数面前人人平等”的表面公正中,不能忽视高考命题方面隐含着的深层次的实质性不公正,因为其直接影响到高等教育机会均等的实现。从命题的过程来看,命题目的确定、考试方式的选择、试题内容(范围和深度)的确定、试卷难度和区分度的把握、题型的设计及试卷的组配、试卷等值等要素都关涉命题公正问题。命题对于高考公正影响最深刻、最广泛,而且一般情况下是隐性的、深层次的,不容易察觉。分析高考命题层面的公正问题和探究命题公正的价值取向,既是研究高考与社会和教育关系无法回避的理论问题,又是直接关乎高考如何改革乃至高考这一“举国大考”命运走向的实践问题。

### 一、高考命题的公正性问题

高考命题有广义和狭义之分。狭义的高考命题是指高考招生出题,并不指试题的组配。如果试题命制采取题库的方式,命题与组配就是分离的,这种方式的命题就是狭义的命题;而如果试题命制不是采取题库的方式,命题和组配就常常联系在一起,这种方式的命题就是广义的命题。国外一些考试机构的试题命制大多采用题库的方式,而我国目前高考试题命制通常是命题和组配结合在一起的。本文讨论的高考命题特指广义的命题。

命题对于考试是根本性的,对于考试公正也是关键性的。美国测试实践联合委员会制定的《教育公平测试实践规范》对命题公正作了较全面的诠释:对参加测试的人应该是公正的,不受年龄、性别、体能、种族、国家、宗教、性倾向、语言背景或其他个人特性的限制。在测试所涉及的诸多

<sup>\*</sup> 本文系全国教育考试“十一五”规划课题“新课程下高考质量评价标准体系及其实施研究”(课题批准号:2006JKS1051)的研究成果之一。

## 高考改革试点方案的定位、分类及结构分析

余澄,王后雄<sup>①</sup>

(华中师范大学考试研究院,湖北武汉430079)

**摘要:** 高考改革试点方案是一个由多变量、多因素构成的复杂系统,是否合理往往取决于这些相关因素与变量之间的协调程度和整合方式。明确高考改革试点方案的定位,了解高考改革试点方案的类型,认识高考改革试点方案的结构,有助于我们回应高考改革实践中的诸多现实问题。

**关键词:** 高考改革试点方案;定位;分类;结构

**中图分类号:**G632.474 **文献标志码:**A **文章编号:**1000-4203(2015)10-0049-07

### College Entrance Examination Reform Program's Location, Classification and Structural Analysis

YU Cheng, WANG Hou-xiong

(Examination Institute, Central China Normal University, Wuhan 430079, China)

**Abstract:** College entrance examination reform program is a multi-variable, complex systems and structures composed of multiple factors, often depends on whether a reasonable degree of coordination and integration of how these factors and variables. Clear positioning of the college entrance examination reform program, to understand the type of college entrance examination reform program, understanding the structure of the college entrance examination reform program will help us respond to the practice of the college entrance examination reform a lot of practical problems.

**Key words:** reform program of the college entrance examination; location; classification; structural

2014年9月,国务院颁发《关于深化考试招生制度改革的实施意见》(以下简称《实施意见》)。之后不久,上海和浙江两地相继发布了高考综合改革实施方案。2015年,李克强总理在两会政府工作报告中强调“让每个人都有机会通过教育改变自身命

运”,表明高考改革进入深水区。深入领会和贯彻《实施意见》,剖析高考改革试点方案的定位、分类及结构,有利于考试招生制度改革的推进和高考制度的完善。

<sup>①</sup> 收稿日期:2015-08-10

基金项目:教育部人文社会科学研究规划基金项目(14YJA880066)

作者简介:余澄(1986—),女,湖北嘉鱼人,华中师范大学考试研究院博士研究生,从事国家治理与考试制度、国际政治研究;王后雄(1962—),男,湖北黄冈人,华中师范大学考试研究院院长,教授,博士生导师,从事教育考试与评价、课程与教学论研究。

## “高考城市化倾向”的成因及矫正

◆王后雄

**摘要:** 高考城市化倾向”其形成有内部和外部多重因素,包括制度性教育的缺陷、资本存量的代际影响、高校招考政策的偏向、高考科目设置和内容取向的偏颇等,但主要还是城乡教育资源分布不均衡的结果。解决 高考城市化倾向”问题应该多管齐下,包括缩小城乡教育资源分布不均衡的状况、完善和调整高校招考政策、建立有效的高考公平制约机制、努力提高考试工作者的公平素养等。

**关键词:** 高考公平; 教育公平; 高考改革; 城市化倾向

随着新课程高考在实验省区的实施,“高考城市化倾向”逐渐成为一个热门话题,什么是 高考城市化倾向”?为何产生“ 高考城市化倾向”?如何应对“ 高考城市化倾向”?本文试图对此做出解析,以期人们客观地看待 高考城市化倾向”问题。

### 一、“高考城市化倾向”形成的外部因素

#### (一) 制度性教育的缺陷

在教育目的上,制度化教育围绕培育符合国家需要的人才而展开。建国后特殊的国际国内环境决定了高等教育高度服从、服务于国家主义目标和发展战略的特征。这种国家本位、服务意识至今依然强烈。在功能方式上,制度化教育是一种英才的筛选体系,如胡森所言:“在具有高度筛选性和(或)竞争性的系统中,不平等依然存在,甚至还在有所加剧……覆盖所有义务教育阶段和单轨制的综合性教育系统……已出现不平等移到前大学阶段或大学阶段的趋势。”<sup>[1]</sup>基础教育领域长期实行的以城镇为核心的重点学校制度,过早地人为性地把学生分流到了不同层次的成长轨道,使大多数农村学生的利益受到了损害。重点学校制度的副产品——择校现象及收费

王后雄/华中师范大学考试科学研究中心主任 教授 (武汉 430079)

制度,打破了单一的分数能力标准,基础教育长期的不均积续,使得参加高等教育入学选拔的学生群体间已经存在着巨大的城乡差距。相对于城市学生,农村学生人力资本存量不足、家庭贫困、教育教学条件匮乏、教师整体素质不高、学校教学方法和手段落后等是造成高考弱势的重要原因,由于贫困,农村学生大多只能在非重点中学的“简易手段”中学习。由于这些学校教学条件及设备、师资力量、教学质量与城市学校相距甚远,必然导致一些个人或群体沦为“差生”。不均等的入学机会,没有可比性的教育条件,必然得不到同样的学业成功。结果弱势学生受到的后天不公平遭遇被偷换为其自身的先天智能劣等。近几年来,我国农村学生与城市学生之间的高考成绩存在着一定的差距,其主要原因是他们的教育资源及教育机会并不相同。

在上世纪90年代中期以来的社会转型阶段,新的社会结构相对定型,出现了各不相同的利益集团,使改革的逻辑和机制发生变化。社会阶层的分化造成了一个包括失地农民、下岗工人、失业和半失业人群等构成的庞大的弱势阶层和底层社会。<sup>[2]</sup>从城市教育投入来看,地方政府可以用足够财力对教育硬

## 高考改革的公平风险分析

余 澄, 王后雄

(华中师范大学 考试研究院, 湖北 武汉 430079)

**摘要:** 高考改革是一项牵涉多方利益的系统工程,经历了 30 多年的演变,其公平的优点和缺点都已充分暴露出来。新一轮高考改革对公平的影响具有复杂性和双面性,而隐藏于其后的是社会、国家和民众对其选择态度的影响。从多视角审视公平,对风险比较大的改革项目造成的公平隐患及冲突进行成因分析,为完善高考改革试点方案、化解高考改革的公平风险提出切实的意见。

**关键词:** 高考改革; 公平分析; 风险评估; 对策

**中图分类号:** G424.74 **文献标志码:** A **文章编号:** 1000-0186(2015)09-0083-07

高考是国家选拔优秀人才的重要方式,也是促进社会流动、实现教育公平的重要工具。对于公民而言,高考公平的一小步,往往就是人生的一大步,成为改变个人命运的天平;对于国家而言,高考改革成败关乎人才强国战略的实施,也是维护社会公平正义、促进民族融合、控制区域发展差距、解决社会矛盾冲突、维系社会稳定的重要手段。随着 2014 年《国务院关于深化考试招生制度改革的实施意见》(以下简称《实施意见》)及相关省市配套文件出台,标志着新一轮高考改革拉开了大幕。上海、浙江作为全国两个高考综合改革试点省份,其“新高考方案”备受社会关注。本文试图通过对新高考方案重点任务的分析和梳理,探讨高考改革项目可能对公平的影响,评估高考改革公平风险性,避免因高考改革而造成公平的隐患及冲突,让高考改革的指挥

棒更好地发挥促进公平、科学选才的正向作用。

### 一、新一轮高考改革的重点任务解析

《实施意见》就深化考试招生制度改革提出了五大重点任务:第一,改进招生计划分配方式。针对区域间高等教育入学机会差距和农村学生上重点高校人数偏少问题,通过宏观调控和专项计划,农村学生进入重点高校的人数明显增加,形成保障农村学生上重点高校的长效机制,提高中西部地区和人口大省高考录取率,进一步促进机会公平。第二,改革考试形式和内容。发挥考试招生在引导素质教育、促进学生全面发展中的重要作用。完善高中学业水平考试,规范高中学生综合素质评价,加快推动高职院校分类考试,深化高考考试内容改革。第三,改革招生录取机制。主要是完善规则,确保招生公平和科学

**基金项目:** 教育部人文社会科学规划基金“国家教育考试公平监测指标体系及保障机制研究”(14YJA880066)

**收稿日期:** 2015-05-11

**作者简介:** 余澄, 1986 年生, 女, 湖北嘉鱼人, 华中师范大学博士生, 主要从事国家治理与考试制度、国际政治研究; 王后雄, 1962 年生, 男, 湖北黄冈人, 华中师范大学考试研究院院长, 教授, 博士生导师, 主要研究方向为教育考试与评价、课程与教学论。

# 基于学科思想方法整合的高考化学命题研究\*

孙建明, 王后雄

(华中师范大学 考试研究院, 湖北 武汉 430079)

**摘要:** 学科思想方法是一个学科的灵魂与核心。界定学科思想方法内涵的基础上, 分析学科思想方法的教育价值并提出基于学科思想方法整合的高考化学命题策略。

**关键词:** 学科思想方法; 高考; 化学命题

**中图分类号:** G633.8 **文献标志码:** A **文章编号:** 1000-0186(2014)03-0067-06

从宏观层面来讲, 一个学科通常是由学科知识、学科能力和学科思想方法三个部分组成, 其中学科思想方法是一个学科的灵魂与核心。

## 一、学科思想方法的内涵

### (一) 学科思想方法概念的界定

简言之, 学科思想就是指对学科发展和学习具有核心价值作用的思想观念。如化学学科中的转化思想、联合思想等。学科方法通常是指一个学科的研究方法、思维方法或是利用该学科知识解决问题的具体途径, 它往往是学科思想的显性化、具体化。如化学学科中的守恒法、终态法等。值得注意的是学科方法和学科思想之间并没有一个明确的界限, 如化学学科中的“结构决定性质”既可以说是一种学科方法, 也可以说是一种学科思想。因此, 学科方法与学科思想可合称为学科思想方法, 特指反映学科本质、思维和规律, 对学科本身及其分支学科的发展有着关键性作用的核心思想观念与方法。<sup>[1]</sup>

### (二) 学科思想方法的分类

有学者依据学科思想方法的层次和外延, 把

学科思想方法分别划分为哲学层次的、一般层次的与具体层次的思想方法以及宏观、中观与微观三种类型的学科思想方法。美国著名学者施瓦布(Joseph J. Schwab)、泰勒(Tyler)等人则把学科思想方法划分为学科组织结构方面的思想方法(discipline organization structure way of thinking)、学科句法结构方面的思想方法(The thinking method of syntactic structures)和学科实质结构方面的思想方法(Disciplines essential structure in way of thinking)<sup>[2]</sup>(见表1)。

表1 施瓦布学科思想方法分类

施瓦布学科思想方法分类	学科思想方法包含的内容及范畴	学科举例(化学)
学科组织结构方面的思想方法	学科类型特征以及与相关学科的逻辑关系, 是普遍的学科思想方法	辩证思想、转化思想、联合思想等
学科句法结构方面的思想方法	学科方法论, 对学科内容的高度抽象与概括, 是具体的学科思想方法	分类与比较、动态平衡与守恒、定性与定量结合、量变与质变、对立与统一、内因与外因等

\* 本文系教育部社科项目“国家考试制度公平标准及评价体系研究”(项目批准号: 985ZB0406)阶段性成果, 并获华中师范大学优秀博士学位论文培育计划资助(项目编号: CCNU13T04027)。

收稿日期: 2013-11-15

作者简介: 孙建明, 1974年生, 男, 安徽淮北人, 华中师范大学考试研究院在读博士研究生; 王后雄, 1962年生, 男, 湖北黄冈人, 华中师范大学考试研究院教授、博士、博士生导师。

# 高考英语测试公平性的多维分析\*

王后雄<sup>1</sup>, 詹先君<sup>2</sup>

(1 华中师范大学 考试研究院, 湖北 武汉 430079; 2 安徽科技学院 外国语学院, 安徽 凤阳 233100)

**摘要:** 借鉴国外学界有关语言测试公平的理论, 结合我国的现实, 对高考英语测试公平性从性别、城乡、阶层、民族、认知风格等多个维度进行了审视, 发现高考英语测试在这些维度上存在程度不等的公平性问题。当前主要从建立高考英语公平性审查制度、加强语言测试公平保障技术的研究、提高语言测试命题者的公平素养等方面实现高考英语测试的公平。

**关键词:** 高考英语; 公平; 多维分析

**中图分类号:** G633.41 **文献标志码:** A **文章编号:** 1000-0186(2011)05-0049-05

英语是我国高校招生的大规模选拔性考试的重点科目, 高考英语对公平性原则进行操作的的第一要务, 就是研究科学合理的选拔标准。然而, 在“分数面前人人平等”的表面公平中, 不能忽略高考英语测试隐含着的深层次的实质性不公平, 其结果直接影响到高等教育机会均等的实现。本文试图借鉴国外语言测试公平研究成果, 检视我国高考英语测试中存在的公平问题, 以期探讨实现高考英语测试公平的基本途径。

## 一、国内外语言测试的公平性研究简述

在21世纪的国际语言测试领域, 语言测试的公平性——特别是大规模语言测试的公平性受到空前的关注, 有关语言测试公平的观点、标准、审查文件不断出台, 专门探讨语言测试公平性的高层次学术会议相继召开。著名的教育测试机构美国教育测试服务中心(Educational Testing Service, ETS)一直走在践行语言测试公平原则的前列。早在1986年, 该机构就规定, 在编制试题的过程中, 除了进行常规的项目分析外, 还必须进行项目功能差异(Differential Item Functioning, DIF)研究, 作为测试效度和公平性的证据之一。进入21世纪, 该机构陆续出台

了包括语言测试在内的考试公平审查的标准和原则, 如2002年的美国教育测试服务中心的质量和公平标准(ETS Standards for Quality and Fairness), 2003年的美国教育测试服务中心公平审查指导(ETS Fairness Review Guidelines), 2004年的美国教育考试服务中心测试的国际性公平审查原则(ETS International Principles for Fairness Review of Assessment)。国际语言测试的权威组织国际语言测试协会(International Language Testing Association, ILTA)也在推动语言测试的公平性方面做出了很多努力, 1997年召开的第19届语言测试研究学术会议的主题就是“语言测试中的公平性”, 会议讨论的内容涉及测试标准、偏差、平等和测试道德等一系列问题。

2000年3月, 国际语言测试协会在加拿大温哥华召开年会, 并通过国际语言测试协会道德准则(Code of Ethics for ILTA)。2005年7月国际语言测试协会在加拿大渥太华举行语言测试研讨会, 专门讨论语言测试的公平性问题, 并通过了国际语言测试协会语言测试实践准则(The Draft ILTA Code of Practice)。许多学者也关注并致力于研究语言测试公平性问题, 早在1990

\* 本文系教育部人文社会科学研究项目“国家考试制度公平标准及评价体系研究”(项目批准号: 09YJA880049)阶段性成果。

收稿日期: 2011-01-15

作者简介: 王后雄(1962—), 男, 湖北黄冈人, 华中师范大学考试研究院院长, 教授, 博士生导师, 主要研究方向: 教育考试与评价、课程与教学论; 詹先君(1972—), 湖北蕲春人, 安徽科技学院外国语学院教师, 主要研究方向: 语言测试与语言教学。

## 新课程下高考地理考试评价标准及 试卷结构技术指标构想\*

王后雄, 汪永鑫

(华中师范大学 考试研究院, 湖北 武汉 430079)

**摘要:**新课程高考地理考试评价标准及试卷结构技术指标构建的依据是《普通高中课程方案(实验)》和《普通高中地理课程标准(实验)》。地理科试卷结构技术指标包括试卷的结构模式、内容要素、目标要素、题型要素、难度要素、分数要素、时限要素、长度要素和等值要素,它是新课程高考地理命题、审题评估监控的标准,是实现试卷及试题质量控制的依据,可以为高考地理试卷及试题质量评价提供系统的可供操作的标准体系和方法手段。

**关键词:**新课程高考;地理;评价标准;技术指标

**中图分类号:**G633.55 **文献标识码:**A **文章编号:**1000-0186(2008)05-0070-06

普通高等学校招生全国统一考试(以下简称“高考”)是由合格的高中生和具有同等学力的考生参加的选拔性考试。高考是我国参考人数最多、社会关注程度极高、影响面极大的常模参照性考试。高考内容的改革是高考改革的重点和难点,也是高考改革的本质方向,它不仅关系到高校的生源质量,还制约着基础教育的改革方向,影响基础教育投资效益以及高中课程的改革与实施。拟订的新课程高考地理考试评价标准及试卷结构技术指标,是高考地理学科质量控制的重要内容,是对高考命题、审题及试卷质量评价监控的标准。建构科学、规范的考试质量评价标准是充分发挥国家大规模教育考试效能的关键。

### 一、高考地理科考试标准的构建

所谓考试标准,是指某类或某种考试对应者在规定测评要素方面所应达到的质量标准及要

求,或谓测评的准则或尺度。根据效度、信度作为衡量考试质量的基本指标,因而可以将影响信度和效度的内容目标、能力层次、题型类型、难度系数、试卷长度、考试时限、测验等值等作为考试质量评价的基本依据。

#### (一) 高考地理科考试标准的确立要素

高考地理科考试标准,应准确反映学科考试内容要素和能力要素的考核要求。依据“课程标准”与“考试大纲”的要求,高中地理以可持续发展观为指导思想,以人地关系为主线,以当前人类面临的人口、资源、环境、发展等问题为重点,以现代科学技术方法为支撑,以培养国民现代文明素质为宗旨;高中地理课程必修三个模块,涵盖了现代地理的基本内容,体现了自然地理、人文地理和区域地理的联系与融合;选修模块涉及地理学的理论、应用、技术各个方面,关注人们生产生活与地理密切相关的领域,突破地

收稿日期:2007-11-22; 修回日期:2008-02-20

\* 本文系全国教育考试“十一五”规划基金课题“新课程下高考质量评价标准体系及实施研究”阶段成果(课题编号:2006JKS1051)。

作者简介:王后雄(1962—),湖北黄冈人,华中师范大学考试科学研究中心主任、教授、博士生导师,主要从事教育考试与评价研究;汪永鑫(1965—),湖北黄冈人,武汉中学地理高级教师,主要从事中学地理教学与研究。

## 中考化学考试标准 及试卷结构技术指标构想\*

王后雄

(华中师范大学 化学教育研究所, 湖北 武汉 430079)

**摘要:**中考化学考试试卷结构建模的技术指标包括试卷的结构模式、内容要素、能力要素、题型要素、难度要素、分数要素、时限要素,它是中考化学命题、审题评估监控的标准,是实现试题及试卷质量控制的依据。将上述试卷结构构建的技术指标,采用中考化学命题多维细目表的方式,可直接用于指导中考化学试题的命题。

**关键词:**中考化学试卷;考试标准;试卷结构;技术指标

**中图分类号:**G633.8 **文献标识码:**B **文章编号:**1000-0186(2007)04-0058-06

《教育部关于基础教育课程改革实验区初中毕业考试与普通高中招生制度改革的指导意见》(以下简称《指导意见》)指出,“考试结果既是衡量学生是否达到毕业标准的主要依据,也是高中阶段学校招生的重要依据之一。”化学考试是中考的重要组成部分,中考化学考试标准及试卷结构技术指标的构建,是中考化学质量控制的重要内容,是对中考命题、审题进行评估监控的标准,是省级教育行政部门落实《指导意见》评估监控制度的重要依据,其科学性、规范性、合理性是考试效能充分发挥的关键。

### 一、中考化学考试标准的构建

所谓考试标准,是指某类或某种考试对应者在规定测评要素方面所达到的水平要求,或谓测评的准则或尺度。中考化学测评要素主要包括内容要素和能力要素两类,水平要求即为不同测评要素的考核标准。构造中考化学考试标准,首先要明确内容要素、能力要素以及各要素的考核标准的构建依据。

#### (一) 中考化学考试标准构建依据

中考化学内容要素、能力要素及其考核标准的确

定,应以初中化学课程标准为基准,并结合高级中等学校教育目标需求、考试目的的需求、初中化学教材内容体系、社会对国民素质的发展需求和考生的身心特征综合考虑。

由于初中化学教学的启蒙性和基础性,掌握化学基础知识既是作为一名合格公民对科学素养的需要,也是继续学习的需要。《全日制义务教育化学课程标准(实验稿)》将初中化学的课程理念定位为提高学生的科学素质。

化学课程标准既体现了基础教育阶段化学学科教育的目标,也充分反映了社会对国民素质的发展需求。中考化学学科在构建考试标准时还应参照高级中等教育培养目标,以高一级学校对入学新生的知识、能力、素质的要求为依据,满足高一级学校人才培养需求。这些构成了中考化学考试目标的内容要素和试题选材的依据。

在考试目的方面,中考化学学科的成绩将与其他科成绩一起作为中等教育择优录取的依据,所以,应有较高的信度、效度、必要的区分度和适当的难度,在考核标准上应当把握。在学科教材内容体系上,应充分考虑包含的知识单元及知识之间的联系和学科间知识的融

\* 本文是全国教育科学“十五”规划重点基金课题“初中升学考试质量控制及评价标准体系研究”成果之一。

收稿日期:2006 11 01

作者简介:王后雄(1963—),湖北黄冈人,华中师范大学化学教育研究所所长,教授,湖北省特级教师,主要从事化学课程与教学论、教育考试与评价研究。

## 高考公平的内涵及属性

◆ 王后雄 李佳 李木洲

[摘要] 就高考公平的本质而言,它是教育领域中人际利益关系的反映、度量与评价,其目的在于实现教育利益分配的公平。只有在教育利益、国家利益、个人利益、个人价值观及社会基础等范围内探讨高考公平,才能解释和揭示高考公平的真正内涵。高考公平是国家根据一定历史阶段发展的需要,在现有社会条件及考试技术的基础上,按照合理性的规范和原则对高等教育入学机会进行分配的一种实然状态。高考公平具有历史性、相对性、主观性、客观性、不充分性、矛盾性、系统性及理想性等属性。

[关键词] 高考公平; 利益分配; 考试技术

[中图分类号] G424.74

[文献标识码] A

[文章编号] 1002-4808(2012)05-0031-05

高考作为“上联高等教育,下引基础教育”的大规模竞争性选拔考试,既关涉人才培养的质量和效果,又关涉高等教育资源分配的公平性和科学性。如果说高考的科学性直接决定其选拔人才的效益,那么高考的公平性则直接影响到民心所向。因此,科学与公平是高考追求的永恒价值。目前,我国各地高考科目设置方案多达二十余种,加之分省命题、自主招生、加分政策等高考相关政策的出台,其导致的直接结果便是试卷等值成为难以克服的问题,考生之间成绩的可比性大大降低。在试卷各异和成绩可比性难以保障的前提下共同分享、分配全国高等教育资源,其公平性必然备受人们的关注。盛赞者有之,诟病者亦不绝。厘定高考公平的内涵与属性,澄清高考公平的本质所在,对于正确理解高考公平与指导高考改革具有重要的理论意义和实践价值。

### 一、高考公平的内涵

就高考公平的本质而言,它是教育领域中人际利益关系的反映、度量与评价,其目的在于实现教育利益分配的公平。由于高考公平隐藏在复杂的社会、教育及考试现象之中,造成人们对高考公平的内涵和性质存在认识上的差异,所以准确界定高考公平的范畴是把握高考公平内涵及本

质的基本前提。对高考公平内涵或本质的理解存在偏差,对高考公平外延的界定明显偏窄,造成人们对高考公平认识上的异化和迷失,这是当今高考公平危机扩大化的一个重要原因,即人们在对高考公平片面理解的基础上谈论高考是否公平。因此,正确理解高考公平的内涵,首先必须准确界定高考公平的内涵范畴。

#### (一) 高考公平的内涵范畴

高考公平作为社会的抽象存在,也具有自身的分界或范畴。如何界定高考公平的范畴,事关能否准确把握高考公平的内涵,因为只有在特定的范畴内讨论高考公平才具有实际意义。笔者认为,高考公平涉及以下几个层面。

第一是个体权利层面。其主要包括考试机会、考试权益及考试资源的分配问题。高考作为社会个体获得高等教育机会的主要途径,高考公平必须体现考生在考试权利上的公平。首先是机会的平等,即必须确保每个社会个体不受家庭背景、种族、性别及资本占有状况等因素的限制与影响,社会的制度安排和非制度安排给每个社会个体的考试机会是平等的。其次是权益的平等,即每个社会个体都享有法律保护的高考权利和高考可能带来的各种利益,例如,考生自主投档选择高校的权利、享受政策规定的加分权利以及自

王后雄/华中师范大学考试研究院院长,教授(湖北武汉 430079);李佳/华中师范大学教师教育学院讲师,博士(湖北武汉 430079);李木洲/厦门大学考试研究中心博士研究生(福建厦门 361005)。

## 从普通高中课程结构变革看高考改革

◆ 王后雄

[摘要] 《普通高中课程方案》的推行,为高考改革、发展既带来了生机,也带来了一些需要慎重处理的问题。课程结构的变革为师生选择性发展、高考科目设置与内容改革、高中学业水平考试推行和高考深层改革等带来机遇,同时也对高考公平与公正性、高考成本、高中选修课程开设、高考的科学性与社会性等产生了一些影响。为此,要充分认识问题的性质及其本质,并从理论研究和实践探索两个方面着手加以解决。

[关键词] 高中; 课程结构; 高考改革; 机遇与困境

[中图分类号] G 632.479

[文献标识码] A

[文章编号] 1002-4808(2008)04-0041-04

2003年4月,教育部推出新一轮《普通高中课程方案》以及高中各科《课程标准》。2004年9月,广东、山东、海南和宁夏4省(区)高一新生开始使用课改教材,普通高中新课程改革由此全面启动。2007年6月,4省(区)率先进行新课程高考。笔者以为:高考改革是一次利益重新调整和分配的重大事件,悠关公平、科学地选才以及课程改革的走向。要保持新课程高考的活力,必须对已实施方案作进一步变革和完善,积极稳妥地进行第二次论证。本文仅从课程结构变革的角度对高考做初步的探讨。

### 一、课程结构变革给高考改革发展带来的机遇

#### (一) 为师生选择性发展提供了保障

我国课程结构长期以来鲜有改革,少有的变革也较多地停留在课程结构的形式化上,而且,由于诸多原因,这些变革在实践中并未得到严格的贯彻和落实。高中新课程虽仍采用必修加选修的形式结构,但模块化的设置方式使选修得以名副其实,从而为师生的选择性发展提供了条件。在新的课程结构体系下,学生能够在必修模块学习的基础上,充分依据自己的志趣、潜能状况以及发展需要选择合适的模块进行学习。这种相对集中的学习可以使学生的知识运用能力和开拓进取精神等得到有效发展。也就是说,多样性、选择化的高中课程结构体系为学生的个性发展、自主选择提供了保障。

不仅如此,新的课程结构还使教师从多方

面、多角度体验和探究教学成为可能,教师可以根据自己的教学潜能以及学科专业的特点选择性地重点研究某些学科和某几个模块课程的教学。这不仅开拓了教师的教学视野,激活了教师的教学研究思路,而且为教师的专业发展提供了多个方向。

#### (二) 为高考科目设置改革提供了条件

可以进行多样化选择的高中新课程结构及其整体性目标要求与现行的单一性高校招生考试存在较大的反差。社会需求的多元化、学生潜能特长的多样化都在新课程体系中有所体现,但两者需要通过高考实现对接。但是,高考的单一性和高招的简单化,使社会需要与学校培养难以实现对接。<sup>[1]</sup>

高中新课程的亮点之一,就是将课程结构分为必修模块和选修模块。问题在于:选修I有100多个模块,这些模块不是所有学校都要开设的,不同学校开设的选修模块乃至同一个学校学生选修的模块都不一样。对选修课的考核是新高考的一个难点。在必修课与选修课的考核中,4省(区)高考设计的改革方案是:<sup>[2]</sup> ① 用科目组的“X”实现基础性与选择性、共同要求与个性发展的统一,以适应高中新课程标准要求;② 在选修模块上采取“超量给题、限量做题”的策略,以方便选修模块开设多的学校和学习不同模块的学生进行选择。

#### (三) 为高考内容改革提供了便利

现代科学正处于迅猛发展过程之中,根据学科的发展成果替换或增减课程内容一直是课

王后雄/华中师范大学考试科学研究中心主任,教授,博士生(湖北武汉 430079)。

## 论教育考试的功能性缺陷与价值冲突

王 后 雄

(华中师范大学考试研究院, 武汉 430079)

**摘 要:** 迄今教育考试在基础教育发挥着其他要素所不可替代的功能。但是, 教育考试自身存在着“利弊兼具”的功能性缺陷, 正反映了它作为一种考试内在的矛盾以及人们对这一问题价值分歧与冲突, 考试恰恰是在动态处理这些矛盾中得到不断发展和完善的。教育考试的目标与方向, 就是要在考试功能正向性与负向性、考试系统内部与外部、考试的社会性与合目的性之间寻求契合与平衡, 发挥教育考试的正面功能, 实现教育考试目的和价值的回归。

**关键词:** 教育考试; 功能; 价值冲突

**中图分类号:** G40 **文献标识码:** A **文章编号:** 1000-5560(2008)01-0047-11

在当今中国的教育领域, 如果论争议最多、矛盾焦点最集中的, 非教育考试莫属。现阶段, 考试理论研究界在有关教育考试的功能问题上都还有不少疑点、难点和盲点。尤其对直接导致考试负面影响、关乎考试存废的考试外部诸因素及其关系的认识, 目前既不深入, 更不全面。这一方面使考试的负面影响难以得到有效克服, 同时由此引发了许多有关考试功过的无味论争, 也招来了不少对考试的无由指罪, 并在我国形成了教育考试的两种最极端的看法, 要么取消考试, 要么对考试进行釜底抽薪的变革。这种“轮回”的表象的背后, 其深处处正反映了教育考试内在矛盾以及人们对这一问题的价值分歧与冲突。

从教育改革的现实来看, 考试作为教育的一种手段能够解决教育中的一些问题, 但是并不能解决教育中的所有问题。对考试的过分迷信和对考试功能的任意夸大, 将使考试肩负过多的期望; 而同时, 考试又常常令人失望, 不可避免地成为了众矢之的。教育考试自产生以来, 尤其是反映教育主体价值观, 体现考试的本质, 事关人及教育发展利害得失的成文法规考试产生以后, 教育考试就少有宁日。正如《中国考试通史》主编杨学为先生所说: “一部考试的历史, 就是争论不断、屡废屡复的历史。”中国自汉魏以降, 西方国家自公元17世纪以后, 上至国家元首, 下至黎民百姓, 有关考试的论争、评说甚至抨击、批判, 可谓代不绝人。争论主要围绕三大问题展开: 一是考试制度的存与废; 二是考试功用的利与弊; 三是考试质量的优与劣。而考试究竟有无存在的社会价值, 则是持久论争的焦点。放眼当今社会,

收稿日期: 2007-10

本文系全国教育考试“十一五”规划课题《新课程下高考质量评价标准体系及其实施研究》基金资助研究成果之一。

**作者简介:** 王后雄(1962-), 男, 湖北黄冈人, 华中师范大学考试科学研究中心主任、教授、博士生, 主要从事教育考试与评价、课程教学研究。

## 国外减缓高考竞争压力的改革及启示

王后雄

(华中师范大学考试研究院, 湖北 武汉 430079)

**[摘 要]** 高考竞争压力已经成为突出的社会问题。许多国家为了减缓高考竞争的压力而进行了一系列改革, 如提供多次考试机会, 为不同考生群体提供分类考试, 增加科目的选择性, 降低高考成绩在录取中的权重, 考试与评价凸显人性化, 改革考试形式和内容等。借鉴其经验, 我国应突出高考制度的稳定性与变革性, 通过改革维系高考的活力。

**[关键词]** 高考; 减缓竞争压力; 改革; 启示

**[中图分类号]** G 473.2

**[文献标识码]** A

**[文章编号]** 1006-7469(2007)11-0060-06

为了应对21世纪激烈的国际竞争, 全球性的教育改革浪潮促使高考在观念、制度与行为层面发生深刻的变化与改革。在一些发达国家中, 对高考的重视程度丝毫不亚于中国, 大学入学考试的竞争同样十分激烈。为此, 政府为了缓解高中教育的应试现象, 在有效利用高考指挥棒作用, 减轻高考竞争压力, 引导基础教育从应试教育向素质教育转变等方面进行了一系列改革, 且取得实效。因此有必要对其进行介绍分析, 并从中寻求有益于改善我国高考制度的启示, 以减缓日益加剧的高考竞争压力, 推进基础教育的改革与发展。

### 一、高考竞争压力已成为突出的社会问题

在当今我国教育领域, 如果说有什么能牵动社会各阶层民众的, 应该说非高考莫属。民意调查显示, 高考已成为“举国大考”。高考关系到千百万青年学生的前途命运和千家万户的切身利益, 承载着广大学生、家长和教师的期望, 高考因而成为典型的高竞争、高利害、高风险的大规模选拔性考试。从表面上看, 高考不过是高等学校招收新生的考试, 但高考又不仅仅是教育考试, 它同时还承担着巨大的社会责任。除了为高校选拔合适人才、以考促学等教育功

能之外, 高考也具有维护社会公平、维护社会稳定、促进并控制社会流动等社会功能<sup>[1]</sup>。社会给高考加载了太多的重负。

从学校教育的导向来看, 由于高考“一考定终身”, 考生的学习也就完全围绕着高考转, 考什么就教什么、学什么。不仅高中的学习成了应试教育, 而且这种应试教育有辐射到初中甚至小学的趋势。学校考试片面地围绕在完成高考升学任务指标上, 从而加剧了学校之间、学生之间的考试竞争。中小學生纷纷争上名校, 其实质就是为将来上好的大学创造条件。近几年, 我国一直强调素质教育, 但并没有取得实质性的突破与进展, 究其原因, 与现行的高校招考制度对基础教育的导向或指挥不无关系。

从高考的竞争过程来看, 高考一直以来是悬在学生和家长头上的一把“利剑”。考生十多年的寒窗苦读, 惧怕因各种难以避免的意外而美梦成空, 也正因为如此, 不少优生及其家长在高考前和高考过程中胆颤心惊, 生怕有点滴闪失和意外, 背负着沉重的心理与精神负担。对学习落后的学生而言, 明知复习无成效, 考不出好成绩, 但还是为形势所逼, 无奈应试。出现的严重后果之一, 就是相当多的学生因此讨厌学习, 甚至讨厌社会和人生。这种将高考作为竞争目标而执着追求和依赖成瘾, 忽视高考结果的价

收稿日期: 2007-07-20

[基金项目] 全国教育考试“十一五”规划课题《新课程下高考质量评价标准体系及其实施研究》(课题编号: 2006JKS1051)。

[作者简介] 王后雄(1962-), 男, 湖北黄冈人, 华中师范大学考试研究院教授, 考试科学研究中心主任, 华中师范大学考试研究院博士生。

# 新课程下高考物理质量评价标准 及试卷结构技术指标构想

■ 王后雄 漆应该

**摘要:** 高考物理质量评价标准的技术指标包括试卷的结构模式、内容要素、能力要素、题型要素、分数要素、长度要素、时限要素、等值要素,它是高考物理学科命题、审题评估监控的标准,是实现试题及试卷质量控制的依据。上述试卷结构构建的技术指标,可直接用于指导高考物理试题的命题,并可作为评价高考物理试卷质量标准的重要依据。

**关键词:** 高考物理;评价标准;试卷结构;技术指标

【中图分类号】G405 【文献标识码】A 【文章编号】1005-8427(2009)07-0018-8

普通高等学校招生统一考试(以下简称“高考”)是一种高利害的大规模教育考试,考试成绩是高等学校招生的主要依据。它对考生家庭、考生所在学校乃至社会都会产生重大的影响。大规模教育考试的目的和性质决定了考试结果误差必须尽可能少,具有高度的可靠性——信度;由于以考试分数为依据对考生进行选拔,因此要求考试结果具有高度的有效性——效度;信度和效度是大规模教育考试最重要的质量标准。虽然我国非常重视大规模教育考试,考试结束后,相关部门也组织有关对考试进行评价,但这种评价主要还是流于形式。到目前为止,仅个别考试机构开始了制定考试的质量标准的工作,整个教育考试界还没有一个公认的考试应该达到的质量标准。

大规模教育考试的信度和效度取决于很多因素,如考试的测量目标;试卷的长度及试卷的结构;考试的试题类型、试题的难度、分数的配置及时限的设计等。物理学科是高考的重要组成部分,高考物理的信度和效度的影响因素是高考命题质量控制与评价的重要内容,是对高考命题、审题进行评估监控的标准,其科学性、规范性、合理性是考试效能充分发挥的关键。

## 1 高考物理考试标准的构建

高考物理考试标准,应准确反映学科考试内容的测量目标,以及学科考试内容要素和能力要素的考核要求。根据普通高等学校对新生知识与技能、方法和能力的要求,依据中华人民共和国教育

# 新课程下高考化学试题 内容选材及能力结构要素分析

王后雄

**摘要:** 新课程下高考内容的改革是高考改革的难点和重点, 关涉高考改革的本质与核心。探讨实施课程标准后新高考化学学科的考试范围和内 容、试题的构造技术及能力要素。试题选材构造技术包括价值取向、学术领域、知识视野、社会背景、教育功能和导向功能等多维要素; 能力要素包括一、二级能力指标, 其中一级指标包括“知识了解”、“知识理解”、“规则运用”和“问题探究”, 在具体命题时还应考虑主体内容的“能力亚层”、“题干特征”及“答案特征”三方面要素。

**关键词:** 新课程; 高考试题; 高考化学试题命题

【中图分类号】G40-058.1 【文献标识码】A 【文章编号】1009-718X(2007)08-0022-07

新课程高考是一项系统、复杂的工程。高考内容的改革是高考改革的难点和重点,也是高考改革的本质方向。它不仅关系到高等教育培养人才的生源质量,还制约着基础教育的办学方向和课程结构的变革,影响基础教育的改革创新及投资效益。化学学科是高考的重要组成部分,高考化学试题结构和内容的技术要素、能力立意建模、试题设计等是高考命题质量控制与评价的重要内容。因此,探讨高考化学命题内容、结构、能力要素与化学新课程的内容关系就成为一个具有理论和实践意义的重要问题。

## 一、高考化学考核目标及内容的确定要素

高考化学考试标准,应准确反映学科考试内容的考核要求。根据普通高等学校对新生文化素质的要求,依据中华人民共和国教育部2003年颁布的《普通高中课程方案(实验)》、《普通高中化学课程标准(实验)》和《2007年普通高等学校招生全国统一考试大纲理科课程标准(实验版)》确定高校理工类招生化学科考核目标与要求,也是高考内容确定及命题的依据。

### (一) 新课程高考化学考核目标与要求

为了有利于选拔具有学习潜能和创新精神的考生,化学科考试以能力测试为主导,并在测试考生进一步

学习所必需的知识、技能与方法的基础上,全面检测考生的化学素养。

化学科命题注重测量学生自主学习的能力,重视理论联系实际,关注与化学有关的科学技术、社会经济和生态环境的协调发展等问题,以促进学生在知识与技能、过程与方法、情感态度和价值观方面的全面发展。

### (二) 新课程高考化学考试范围和内 容

根据普通高等学校对新生科学素养的要求,按照既保证与全国普通高考招生统一考试的要求基本一致,又有利于实验省区实施普通高中化学课程标准的原则,参照《普通高中化学课程标准(实验)》,高考化学科考试范围可分为必考内容和选考内容(见表1)。

新的高考大纲与高中新课程改革进行了有效的衔接,但在实际中也显露出一些不可回避的问题,与高中新课程教学中的大量选修模块课程相比较,五省区新高考方案中,每门试卷中所涉及的选修模块的比重偏小,限制过死,体现多样性、选择性的分值也不多,这样可能导致高中一些选修课程多样性的缺失,导致师生对所考化学学科内容的再次选择(如师生只会选择少数固定的模块学习,不考的内容会放弃),因此必须在高考实践中找到一个“平衡点”,力求在为高等学校选拔合格的新生、实现新课程改革目标和减轻学生的过重学习负担三者之间实现科学的统一。

王后雄 华中师范大学考试研究院 教授 430079

# 弱势群体教育考试公平： 问题、成因及对策

■ 余澄 王后雄

**摘要：**弱势群体教育考试公平是教育公平的重要组成部分，教育考试公平权是弱势群体应该拥有的重要权利之一。目前弱势群体高等教育入学机会存在较大差距，经济落后与贫困环境是决定性原因，区域教育资源配置失衡是基础性原因，弱势群体救助政策失真根本性原因，家庭出身和生理因素制约是内生性原因。从根本上遏制转轨中弱势群体教育考试公平差距，政府应承担起对弱势群体教育支持的主要责任，建立弱势群体教育救助制度和救助管理体系，从“弱势补偿”战略高度完善和调整考试招生政策，构建弱势群体教育支持和考试公平保障机制。

**关键词：**弱势群体；教育考试公平；弱势补偿

【中图分类号】G405 【文献标识码】A 【文章编号】1005-8427(2016)05-0036-12

高考是一种高竞争、高利害、高风险的大规模国家选拔人才的重要方式，是社会公平公正的底线，关系到千家万户的切身利益。恢复高考30多年来，高考改革的步伐历经曲折，而公平一直是影响和制约高考改革的重要因素。相对于其他阶层，高考是弱势阶层民众改变命运、实现社会流动的最重要乃至唯一的途径。目前关于弱势群体的考试公平评价观点和改革意见建议缺乏，显然不利于高考改革的科学决策和教育公平的发展。因此，研究弱势群体考试公平的问题、成因及对策，对于推进新一轮高考改革和促进社会公平公正发展具有十分重要和迫切的现实意义。

## 1 弱势群体与弱势群体教育考试公平概述

弱势群体是指由于各种原因而处于不利的社

会地位的人群或阶层。我国的弱势群体主要是由农村地区与西部地区的贫困人口、妇女、少数民族、残疾人、下岗工人及其子女组成的。<sup>[1]</sup>弱势群体成为我国一个独特而带有普遍性的社会问题，也是当今国际社会共同关注的社会公平发展的焦点问题。弱势群体教育可分为弱势群体本身的教育和其子女或有特殊需要的儿童、未成年人的教育两类。研究弱势群体教育考试公平特指第二类的教育群体在获得高等教育入学机会的状况。在教育考试竞争过程中，由于群体的力量、权利、资源、能力相对较弱，因而在分配、获取高等教育（尤其是优质高等教育）机会方面较少较难。在我国，效率优先、优胜劣汰的竞争机制引入到高考选择过程中，于是那些经济收入、社会地位、生存状况、生理特征、体能状态和分配能力较弱（较差）的人逐步落入

本文系教育部人文社会科学研究规划基金“国家教育考试公平监测指标体系及保障机制研究”（项目批准号：14YJA880066）的阶段性研究成果。

【作者简介】余澄，女，华中师范大学，在读博士（武汉 430079）

王后雄，男，华中师范大学考试研究院院长，教授，博士生导师（武汉 430079）

# 国家教育考试公平监测指标体系的 缺陷及对策研究

■ 余澄 王后雄

**摘要：**缺少有效的外部质量证据是国家教育考试公平监测指标体系的缺陷。这不仅制约了招考公平质量的提高，而且影响了招考公平的权威性与公信力。问题成因是对教育考试服务的本质属性缺乏认识。解决该问题，需要从服务理念入手，以“服务对象”的角度审视招考公平，完善教育考试公平监测指标体系，提高招考公平监测的正效应。

**关键词：**国家教育考试；公平；监测指标体系；缺陷及对策

【中图分类号】G405 【文献标识码】A 【文章编号】1005-8427(2015)06-0003-5

国家教育考试的公平不仅直接关系到考试本身的信誉和质量，关系到公共利益社会秩序和公民的切身利益，同时也关系到建设人力资源强国目标的实施，也直接影响到社会的稳定。当前，在教育领域存在一些社会质疑的招考公平问题，这些质疑、批评严重影响了公众对国家教育考试公平的信任度。然而，如何构建公平监测指标体系？公平监测证据是否科学？公平监测权威性及公信力如何？这些均成为招考公平亟待解决的现实问题。构建我国教育考试公平监测指标体系既是解决国家教育考试公平的学理性当务之急，也是促进社会公平正义、实现人力资源强国之现实需要。鉴于此，本文将对国家教育考试公平监测体系的缺陷进行分析，并针对存在的问题提出相应对策。

## 1 国家教育考试公平监测指标体系的缺陷分析

高考是我国大众普遍认可、最符合国情、最体

现公平的国家教育考试制度。高考的本质属性是服务于人，服务于教育，服务于高校选拔人才，服务于国家公正有效分配各类人力资源。在我国，由于历史和社会原因，高考由国家主办，依靠行政行为实施。高考实际在借国家和政府公信力举办，故无论公平呼声与质疑如何，考试机构都可以不加理会。高考的服务本质之所以长期漠视，原因在于长期处于“卖方市场”、“公信寻租”和“行政垄断”。考生处于从属“奴性”地位，只有听从“公平”任意分配。久而久之，这种不合理被默认，成为了一种习惯。

随着社会发展，广大民众对考试公平的要求不断加强，提高中西部地区和人口大省高考录取率、增加农村学生上重点高校人数的呼声强烈，高考的公平问题凸显出来，并由此延伸到考试公信力、国家公信力的层面。故国家从建设人力资源强国和教育公平的战略高度提出，“着力完善规则，确保公正公平；体现科学高教，提高选拔水平”，解决区域、城乡、阶层入学机会存在的差距。此次国务院《关

本文系教育部人文社会科学研究规划基金“国家教育考试公平监测指标体系及保障机制研究”（项目批准号：14YJA880066）的阶段性研究成果。

【作者简介】余澄，女，华中师范大学博士生（武汉 430079）

王后雄，男，华中师范大学考试研究院，院长，教授，博士生导师（武汉 430079）

## 新课程高考化学试题命制与 课程标准一致性研究

■ 王后雄 孙建明

**摘要:** SEC作为一种“一致性”定量分析的方法,具有诸多优点,美国联邦政府和十几个地方州市都采用这种方法进行考试命题与课程标准一致性的评价研究。通过采用SEC定量分析的方法,从内容要素、认知层次两个维度2010年、2011年和2012年(新课标)全国高考理综卷的化学试题与课程标准,发现它们之间不存在统计学范畴内的显著一致性。内容要素分布上,三个年份的试题与课程标准存在着明显的差异;认知层次分布上,三个年份的化学试题与课程标准的要求基本一致。

**关键词:** 高考化学试题;课程标准;一致性研究

【中国分类号】 G405 【文献标识码】 A 【文章编号】 1005-8427(2013)11-0010-12

国家化学课程标准是高考化学试题命制的依据,<sup>[1]</sup>高考化学试题命制对化学新课程的实施又有着重要的导向和反拨的功能。明确国家化学课程标准对高考化学试题命制“质”的规定性,有利于强化教师基于课程标准的教学动机,有效把握化学常识教学的深广度,大幅度提升化学基础教育水平。同时,高考化学试题命制与国家化学课程标准保持一致性又是极其重要的,不仅直接关系到高考化学试题命制的信度与效度,而且直接影响到课程预设和课程实施间的吻合度。<sup>[2]</sup>基于以上原因,我们采

用SEC(Surveys of Enacted Curriculum)定量一致性分析的方法,从内容要素维度和认知层次维度两个方面来研究高考化学试题与课程标准之间的一致性,以期为新课程背景下的高考化学试题命制提供可资借鉴的实证依据。

### 1 研究方法

#### 1.1 研究样本的确定

(新课标)全国高考理综化学卷是由教育部考试中心组织命制,具有一定的代表性和权威性。故

本文系教育部华中师范大学国家教育创新平台理论创新研究项目(项目批准号:985ZB0406)的研究成果之一;湖北省教育科学“十二五”规划重点课题“中小学教师教学诊断能力发展机制研究”(课题编号:2012A005)阶段性成果。本文获华中师范大学优秀博士学位论文培育计划资助(Excellent doctoral dissertation cultivation grant from Huazhong Normal University)

【作者简介】 王后雄,男,华中师范大学化学教育研究所,所长,教授,博士生导师(武汉 430079)

孙建明,男,华中师范大学化学教育研究所,在读博士研究生(武汉 430079)

## 新课程下高考质量与 新课程改革发展关系探析

■ 王后雄 李木洲

**摘要:** 新课程改革是推进我国素质教育的重要举措,而新课程下高质量的高考是推行新课程改革的有力保障,同时,科学合理的新课程改革也为新课程高考改革提供良好的基础和发展空间。制约高考质量的因素主要来自两个方面:一是内在的,包括新课程目标和新课程标准体系;二是外在的,包括社会、经济、政治及考生身心发展规律等。制约新课程改革的因素主要来自新课程高考的科目设置、内容选择、评价及录取方式等的变化。

**关键词:** 高考质量;新课程改革;关系

【中国分类号】 G405 【文献标识码】 A 【文章编号】 1005-8427(2010)09-0003-10

为全面推进素质教育,自2001年6月起,教育部相继颁布了《基础教育课程改革纲要》、《普通高中课程改革方案》等一系列课程改革文件。2004年秋,山东、广东、海南、宁夏4省区首先承担了新课改实验任务,并于2007年顺利进行了新课程高考。截至2009年,全国已有21个省市区成为新课改实验区,而开始实施新课程高考方案的省、市、自治区已达11个之多。高考作为我国现阶段最为重要的教育考试之一,肩负着为高等教育选拔人才、评价基础教育质量及分配高等教育资源的职能,尤其对我国基础教育课程改革具有极其重要的反控作用。而高考的质量如何,则直接关系到上述职能能否实现

及其实现的程度。因此,研究高考质量与新课程改革发展的关系,对控制未来高考质量、构建科学的教育评价标准体系以及全面推进素质教育,具有极为重要的现实意义。

### 1 高考质量的内涵

“质量”是一个内涵十分丰富的概念,在物理学上它表示“物体所含物质的多少”或“描述物体的惯性的物理量”;在ISO质量体系系统中它是指“一组固有特性满足要求的程度,也可以看作是产品和服务满足顾客需求的能力”。前者强调物质形态的量,后者则强调非物质形态的满足服务对象的程度。而高

本文为全国教育考试“十一五”科研规划课题“新课程下高考质量评价标准体系及实施研究”(2006JKS1051)阶段成果。

# 新课程下高考生物考试评价标准与试卷结构技术指标构想

■ 王后雄 徐启发

**摘要:** 新课程下高考生物评价标准及试卷结构技术指标构想的依据是《基础教育课程改革纲要(试行)》的基本精神和《普通高中生物课程标准(实验)》。高考生物试卷结构建模的技术指标包括试卷的结构模式、内容要素、能力要素、题型要素、难度要素、分数要素、时限要素,它是高考生物命题、审题评估监控的标准,是实施试题及试卷质量控制的依据。将上述技术指标以多维细目表的方式列出,可直接用于命制高考生物试卷。

**关键词:** 新课程; 高考生物; 评价标准; 技术指标

【中图分类号】G405 【文献标识码】B 【文章编号】1005-8427(2008)07-0017-7

## 1 问题的提出

自2004年以来,普通高中课程改革方案在山东、广东、河南、宁夏4省区开始试验实施。2007年4省区共有120万考生参加新课程高考。随着课程标准的推广,将有更多省市更多考生参加新高考。由于各地执行课程标准时选用的教材版本不同(生物新课标教材有人教版、苏教版、地图版、浙江科技版、河北少儿版等),分省命题执行的高考考查方案的差异,导致各地对课程标准的评价标准各有差异,课程标准的执行与高考试题命制的矛盾更加突出,存在着标准不一、质量失控等问题,严重影响了新课程高考效能的发挥。高考生物考试评价标准及试卷结构技术指标的构造,是高考生物科质量控制的重要内容,其科学性、规范性、合理性是充分发挥考试效能的关键。

## 2 新课程下高考生物考试评价标准的构建

### 2.1 高考生物考试在教学评价中的地位

评价是实施素质教育、促进学生主动发展的一种手段。评价应以《普通高中生物课程标准(实验)》(以下简称《课程标准》)为依据,根据课程目标和具体的教学目标进行,要客观、公正、合理。下面给出了高中生物课程标准评价的框架图解(见图1)。

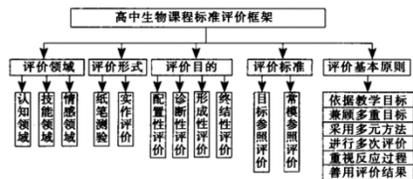


图1 高中生物课程标准评价体系

# 中国高考改革40年艰难探索历程及价值追求

王后雄<sup>1</sup> 詹先君<sup>2</sup>

(1. 华中师范大学, 湖北 武汉 430079; 2. 安徽科技学院, 安徽 滁州 233100)

**摘要:** 在高考恢复的40年间, 高考改革既有技术方面的改进, 也有制度方面的革新, 体现了多元价值追求。这段探索历程留给后人充满智慧的经验 and 无比宝贵的借鉴, 为将来高考的提升拓展了无数新的视角, 未来的高考将会更好地指导中学教学, 更有利于高校的选才。

**关键词:** 恢复高考; 高考改革; 高考

【中图分类号】G405 【文献标识码】A 【文章编号】1005-8427(2017)01-0037-6

DOI: 10.19360/j.cnki.11-3303/g4.2017.01.007

狭义而言, 高考是国家为了全国高等学校招收新生而举行的考试; 广义而言, 高考既是高等学校招生考试, 也指与其相伴的录取制度、政策和法规。这里要探讨的高考改革中的“高考”含义, 是指广义的高考。1977年, 在时任国家副总理邓小平同志的主持下, 高考制度恢复, 千百万人的命运因此而改变, 恢复高考不仅挽救了中国教育, 也挽救了整个中国。然而, 随着时间的推移, 面临“文革”后的新环境、新条件、新技术、新思想, 高考在实施过程中的问题不断显露, 为了使高考适应社会的不断发展, 突破有关考试的认识、环境和功能等方面的局限, 考试主体和客体不懈努力, 不断地对其进行改革完善。高考恢复的40年间, 30多次大大小小的改革, 既有技术方面的改进, 也有制度方面的革新, 体现了多元价值追求, 映射了中国考试人的不懈追求。值此高考恢复40周年之际, 有必要对高考

改革探索历程及其价值追求进行回望, 为未来的高考改革提供启示和借鉴。

## 1 以人为本的改革

以人为本是40年高考改革的首要价值追求。廖平胜先生指出: “考试是人存在和发展的一种方式。考试是以为人为主、客体的对象性活动。它反映了人与自然、人与事、人与人的关系, 具有知识性、依存性和目的性。同时, 人参与考试的过程, 并非仅仅是被动接受检测的过程, 而且是能动展示自我的过程, 立志进取的过程, 与人交互影响的过程, 培养综合适应能力的过程, 是人认识自我、展现自我、发展自我、实现自我的基本方式和手段之一。”<sup>[1]</sup>正因为如此, 作为认识自我、展现自我、发展自我、实现自我的高考, 必须个性化、差异化、精细化。正是从以人为本出发, 我国的高考改革进行了分省命

本文系教育部人文社会科学研究规划基金“国家教育考试公平监测指标体系及保障机制研究”(批准号: 14YJA880066)的阶段性成果。

【作者简介】 王后雄(1962—), 男, 华中师范大学考试研究院院长, 教授, 博士生导师;  
詹先君(1972—), 男, 安徽科技学院外国语学院, 副教授。

# 高考理科定量化试题“静态难度”影响因素的分析及调控

皇甫倩<sup>1</sup> 王后雄<sup>2</sup>

(1. 西南大学, 重庆 400715; 2. 华中师范大学, 武汉 430079)

**摘要:** 对试题难度的调控是提高高考命题质量的关键。定量化试题, 作为高考理科综合考试中分布范围最广的一类试题, 其难度将会直接影响到试卷的质量。本研究在前人的实践经验和研究结论的基础上, 以近10年部分高考理科定量化试题为研究对象, 科学分析影响高考理科定量化试题“静态难度”的9种因素, 并据此提出相应的调控措施, 旨在为试题命制提供有意义的参考。

**关键词:** 高考; 高考命题; 定量化试题; 静态难度; 命题技术

【中图分类号】G405 【文献标识码】A 【文章编号】1005-8427(2017)05-0039-8  
DOI: 10.19360/j.cnki.11-3303/g4.2017.05.007

高考作为一种重要的选拔性考试, 其试题的难度一直都是社会关注的焦点, 它既是评价试题和考试效果的重要质量指标, 也是试题命制工作中进行质量控制的主要内容, 它在很大程度上决定着考试目标的实现程度<sup>[1]</sup>。定量化作为理科的重要特征, 是高考考查学生的重要能力之一。纵观近年高考命题趋势, 定量化试题考查的范围逐年增大, 定量化试题在考查学生科学素养中的作用日益突出, 这成为考生较难把握、较易失分的题型。由此可见, 定量化试题不仅影响着试题的难度, 也影响着人才的选拔。一般我们所探讨的试题难度指的是“相对难度”, 与考生群体的水平有关, 它主要是以通过的人数与全体参试人数的比值或全体参试人数的平均得分与该题满分的比值来评判的<sup>[2]</sup>。这种难度数值极易受外界环境的影响, 稳定性较差。因此, 研究试题的静态难度, 并探究其影响因素是必要的, 而且是有价值的。

【作者简介】皇甫倩(1987—), 女, 西南大学化学化工学院, 讲师;  
王后雄(1962—), 男, 华中师范大学教师教育学院副院长, 教授, 博士生导师。

## 1 “静态难度”的内涵阐释

当前学术界对试题“静态难度”鲜有涉及, 并未界定出统一明确的说法, 本研究将通过对相关文献的梳理, 结合研究重点给予初步的界定。

### 1.1 “静态难度”的界定

当前对“静态难度”的界定主要有以下几种观点:

能力说: 是指考试内容要求学生在解题过程中所需要具有的心智活动水平能力层次的高低, 它反映了试题与学科知识、能力要求的适应程度<sup>[3]</sup>。

属性说: 静态难度是由问题本身的特征属性决定的<sup>[4]</sup>。

要素说: 能力因素、广度因素(即知识点的多少)、深度因素(即考核知识内容深浅的程度)是构成静态难度的3种要素<sup>[5]</sup>。

通过对上述观点的整理和分析, 本研究对试题

# 基于核心素养提升的教育考试策略研究

孙建明<sup>1,3</sup> 王后雄<sup>2</sup> 王祖浩<sup>3</sup> 汪峰<sup>4</sup>

(1. 湖北第二师范学院, 武汉 430205; 2. 华中师范大学, 武汉 430079;  
3. 华东师范大学, 上海 200062; 4. 安徽省淮北市教研室, 安徽淮北 235000)

**摘要:** 核心素养主要指学生应具备的能够适应终身发展和社会发展需要的必备品格和关键能力。落实发展学生核心素养需要从整体上推动各教育环节的变革。通过教育考试策略研究, 探讨如何通过教育评价落实核心素养的教育战略性问题。

**关键词:** 核心素养; 教育考试; 核心素养测评; 教育评价

【中图分类号】G405 【文献标识码】A 【文章编号】1005-8427(2017)08-0026-6  
DOI: 10.19360/j.cnki.11-3303/g4.2017.08.005

2016年9月13日, 《中国学生发展核心素养》正式颁布, 该成果由教育部委托北京师范大学, 联合国内上百位知名专家历时3年完成。落实发展学生核心素养需要从整体上推动各教育环节的变革, 最终形成以学生发展为核心的完整育人体系。具体而言, 主要有三个方面的落实途径: 一是通过课程改革落实核心素养; 二是通过教学实践落实核心素养; 三是通过教育评价落实核心素养。美国教育心理学创始人桑代克(E. L. Thorndike)指出: “凡客观存在的事物都有其数量。”<sup>[1]</sup>著名教育测量专家麦柯尔(W. A. Mccall)博士更进一步指出: “凡有数量的东西都可以测量。”<sup>[2]</sup>纵观国内外学者的相关研究, 前两个方面的研究成果较多, 而第三个方面的研究面临诸多挑战, 亟待教育研究者作出正面回应。

## 1 核心素养的内涵

20世纪90年代以来, 国际经济合作与发展组织(OECD)、联合国教科文组织(UNESCO)、欧洲联盟(EU)等国际组织以及美、英、日、新加坡、中国台湾等国家和地区积极开展学生核心素养的研究。2014年, 教育部颁布的《关于全面深化课程改革, 落实立德树人根本任务的意见》明确提出将研制与构建“学生发展核心素养体系”, 以落实党的十八大提出的“立德树人”总体育人目标。核心素养以培养“全面发展的人”为核心, 分为文化基础、自主发展、社会参与3个方面, 综合表现为人文底蕴、科学精神、学会学习、健康生活、责任担当、实践创新6大素养, 具体细化为国家认同等18个基本要点。

本文系全国教育科学“十二五”规划教育部重点项目“基于全国统一命题趋势下高考试题命制公平监测体系构建与实施研究”(项目批准号: DJA150289)和教育部人文社会科学重点研究基地重大项目“基于核心素养的课程标准研制的阶段性成果, 并获中国博士后科学基金第58批面上资助(资助编号: 2015M580301)。

【作者简介】孙建明(1974—), 男, 湖北第二师范学院, 讲师, 华东师范大学课程研究所博士后;  
王后雄(1962—), 男, 华中师范大学考试研究院, 教授, 博士生导师;  
王祖浩(1958—), 男, 华东师范大学, 教授, 博士生导师;  
汪峰(1963—), 男, 安徽省淮北市教研室, 教研员, 特级教师。

# 高考改革:刚性条件的坚守是 确保教育公平的门槛条件

——兼论素质教育与应试教育的应然与实然的关系

王后雄 邓阳

(华中师范大学考试研究院,武汉 430079)

**摘要:**全面推进素质教育是我国教育改革的战略。然而,在当前的教育实践中,素质教育主导下推出的一系列高考改革政策却伴随着教育不公平因素的产生。高考综合改革的“综合评价”“多元录取”一旦脱离统一考试刚性录取条件,公平性将遭受重创,科学性也将沦为空谈。应试教育通过高考证明其存在的合理性,又以高考的刚性与精细选拔成为实现教育公平的门槛条件。素质教育与应试教育的应然与实然的背离是客观存在、长期存在的问题,这是中国现今国情诸多因素交互作用的结果,协调好两者的应然与实然之间的关系,在应然与实然间找寻一种有效的互认机制,才是缩短梦想与现实之间守望距离的应有之义。

**关键词:** 素质教育;应试教育;高考;教育公平

【中图分类号】G405 【文献标识码】A 【文章编号】1005-8427(2017)09-0030-6

DOI: 10.19360/j.cnki.11-3303/g4.2017.09.004

我国于20世纪末提出素质教育的理念,并从国家层面提出了诸多战略性文件。例如,1999年6月,国务院颁布《中共中央国务院关于深化教育改革,全面推进素质教育的决定》(以下简称《决定》),明确强调了教育在我国综合国力的形成中所处的基础地位,而劳动者的素质决定了国家的强弱。随着素质教育的推进,作为教育评价重要手段之一的考试被推上了风口浪尖,成为许多人诟病素质教育推行不利的由头。但是,伴随着对应试教育、考试的片面认识,新时期高考改革政策的出台,一系列教育公平问题也滋生出来。为什么应试教育有人视为狗血,有人捧为鸡汤?为什么素质教育与应试教育的正确关系总是难以辨清?为什么国家大力推

行素质教育,而应试教育却在现实大行其道?为什么高等教育入学机会普遍增加,而民间对“寒门再难出贵子”的公平呼声越来越高?……中国恢复高考40年,这些追问仍令我们深思。

## 1 素质教育背后的教育公平反思

我国实施的素质教育在很大程度上是国家基于马克思关于人的全面发展学说,综合我国国情作出的决定。20余年素质教育的发展也证明了此决策的正确性。

马克思主义关于人的全面发展的学说是我国教育目的的理论基础<sup>[1]</sup>。培养学生的全面发展是我国实施素质教育的内容与途径,正如马克思认为的

本文系教育部人文社会科学研究规划基金项目“国家教育考试公平监测指标体系及保障机制研究”(项目批准号:14YJA880066)的研究成果。

【作者简介】王后雄(1962—),男,华中师范大学考试研究院院长,教授,博士生导师;  
邓阳(1988—),男,华中师范大学考试研究院讲师,博士。

# 基于“素养为本”的 高中化学学业水平考试命题研究

王后雄

(华中师范大学考试研究院,武汉 430079)

**摘要:**化学学科核心素养包括宏观辨识与微观探析、变化观念与平衡思想、证据推理与模型认知、实验探究与创新意识、科学态度与社会责任5个维度。这5个维度既体现了具有化学学科特质的思想和方法,也从科学探究与创新意识的实践层面激励了创新的科学实践活动,并实现了科学态度与社会责任更高层次的价值追求。阐述这些维度的化学的含义和价值,并使之具象化,为基于“素养为本”的高中化学学业水平考试命题指明了方向。高中化学学业水平考试既要考查5个素养的不同层次,又要以实际问题为测试任务,以真实情境为测试载体,以化学必备知识、思想方法和核心观念为解决问题的工具,运用多种策略编制试题,以充分利用好“考试指挥棒”的作用,为发展学生的化学核心素养这一化学课程的根本目的服务。

**关键词:** 高考改革;学科核心素养;学业水平考试;命题;化学

【中图分类号】G405 【文献标识码】A 【文章编号】1005-8427(2018)01-0027-12

DOI: 10.19360/j.cnki.11-3303/g4.2018.01.005

2014年3月30日发布,教育部颁布的《关于全面深化课程改革 落实立德树人根本任务的意见》(以下简称《意见》)明确提出,将“发展学生核心素养体系”的研制与构建作为推进课程改革、深化发展的关键环节<sup>[1]</sup>。《意见》的出台为课程改革指明了方向,即落实立德树人是课程改革的核心和根本任务,发展学生的核心素养是育人的重要目标。2014年12月,教育部发布《关于普通高中学业水平考试的实施意见》,对学业水平考试提出具体的指导性意见,指出根据国家发布的普通高中课程方案和课程标准的规定及要求确定考试范围和内容;规定命题应紧密联系社会实际与学生生活经验,在全面考查“双基”的基础上,注重加强对能力的考查;规定自主选考科目的学业水平考试计入高校招生录取总成绩;学业水平考试成绩合格,作为普通高中学业毕业及高中同等学力认定的主要依据<sup>[2]</sup>。两份文件

均明确指出,不仅课程标准需要基于核心素养来设计和实施,学生的学业评价标准和学业水平考试标准同样需要围绕核心素养进行设计。

2016年9月公布的“中国学生发展核心素养”提出要发展学生6大核心素养<sup>[3]</sup>。基于此,化学课程标准、化学学业评价标准、化学学业考试标准都需要基于化学核心素养来设计和实施。那么,应该如何构建基于化学学科核心素养的学生水平评价指标,怎样对学生的化学学科核心素养发展水平进行科学有效的测评?这无疑成为教育测量领域面临的一个非常富有挑战性的难题。

## 1 化学学科核心素养的界定及要素分析

### 1.1 化学学科核心素养概念的界定

基于核心素养的定义,化学学科核心素养是指学生在接受化学教育过程中逐步形成的适应个人终

本文系教育部人文社会科学研究规划基金“国家教育考试公平监测指标体系及保障机制研究”(项目批准号:14YJA880066)的阶段性研究成果。

【作者简介】王后雄(1962—),男,华中师范大学考试研究院院长,教授,博士生导师。

## 评美国高校招生改革“新模式”

王后雄

(华中师范大学考试研究院, 武汉 430079)

**摘要:**美国高校招生“新模式”是美国权力精英利用招生改革的话语权,以“科学选才”之名阶层既得利益站台,以获得更多的名校教育机会。这种标榜为“科学”的改革,必将成为精英阶层固化的凝固剂,削弱社会的代际流动程度,损害代际公平。我国高考改革不应复制美国模式,应在全球化改革浪潮中作出适合我国国情的近切选择。针对我国目前正在实施的“三位一体”综合评价,应合理控制招生规模,坚守高考录取的刚性条件,加大对弱势群体补偿和扶持力度。高考改革对“公平性”与“科学性”的双重坚守才能让社会更有温度。

**关键词:**高校招生;高考改革;能力档案;综合评价

【中图分类号】G405 【文献标识码】A 【文章编号】1005-8427(2018)03-0011-4  
DOI: 10.19360/j.cnki.11-3303/g4.2018.03.004

伴随着全球化的进程,世界各国不约而同地启动了旨在满足人才选拔国家战略需求的系统性高考综合改革,采取了一些十分相似的改革理念和举措。向来成为国际教育浪潮领头羊的美国教育界正掀起一项颠覆性的、全新的评价体系改革与讨论,“唯分数论”将不复存在。一项新评价模型(A New Model)——“能力档案”未来有可能成为全美大学录取新生的评价依据。

据报道,2017年5月10日,由逾百所美国顶尖私立高中组成的MTC联盟(Mastery Transcript Consortium)推出的A New Model评估系统,目前已获得美国大学申请系统Coalition for Access, Affordability and Success(CAAS)的支持,使用CAAS申请系统的学校包括哈佛大学、耶鲁大学、普林斯顿大学、哥伦比亚大学、斯坦福大学等80余所美国著名高校。

本文系教育部人文社会科学研究规划基金项目“国家教育考试公平监测指标体系及保障机制研究”(项目批准号:14YJA880066)的研究成果。

【作者简介】王后雄(1962—),男,华中师范大学考试研究院院长,教授,博士生导师。

A New Model抛弃了传统的A-F评分制,专注考查学生的8项能力,具体包括:分析和创造力、沟通能力、领导力及团队合作能力、信息技术及数理能力、全球视野、适应性与探索力、诚信与决策力、心智习惯8项能力。这8项能力又被细分为61项更加具体的子能力<sup>①</sup>。

相较美国传统的评分制,新的能力评估系统有三大特征:一是不看成绩,突出能力导向。学生每修得一门课程,将获得相应的能力学分(Mastery Credits)。这种能力学分呈现的不是A-F等级分,而是“已经可以掌握和使用高等数学”“可以理解并熟练使用两种语言”这样的质性评价。二是动态记录,一目了然。MTC联盟为A New Model开发了一个技术平台,实现电子记录,实时更新学生的能力学分,并提供学生的课堂作品、论文、视频展示等材

## 高考改革试点方案的多维评估方法研究

余澄 王后雄

**【摘要】**新一轮高考改革是我国恢复高考以来最为全面、深刻、系统的一次制度改革。为了避免理论决策的误判,高考改革方案的可行论证和实施成效尚需实践检验。在试点过程中开展高考改革方案的评估是高考改革的一个重要环节,如何创建科学可行的多维度评估方法,则是对高考改革决策者和试验者智慧和能力的考验。高考改革试点方案的成效不仅仅是结果的问题,它涉及到方案的制定、执行、落实(或享受)、监控、反馈、终结等多个环节。只有从全过程、多维度来评估高考改革试点方案的成效,才能够发现改革过程各阶段与改革结果之间的因果关系,才能够科学而全面地考察高考改革试点方案的绩效和成败,从而为完善高考改革试点方案提出更具针对性的改进、完善措施。

**【关键词】**高考改革方案;试点;多维度;评估方法

【中图分类号】G632.474 【文献标识码】A 【DOI编码】10.16518/j.cnki.ema.2015.08.009

2014年9月,《国务院关于深化考试招生制度改革的实施意见》正式发布,就深化考试招生制度改革作出了系统设计和全面部署。随后上海、浙江两地出台了高考综合改革试点方案,正式拉开了新一轮考试招生改革的大幕。然而,高考改革方案尚属理论层面的“满意答卷”,实践才是检验高考改革成败的唯一标准。可以毫不夸张地说,新一轮高考改革能否担当起当代使命,能否得到主流民意的认可,能否避免理论决策的误判,尚需在试点中开展跟踪评估。本文试图在对国内外考试招生改革政策评估过程的分析和梳理基础上,探求高考改革试点方案的多维度评估方法。

对高考改革试点方案的过程进行评估,应从高考改革试点方案的制定、执行、落实(或享受)、监控、反馈、终结六个阶段进行。<sup>①</sup>高考改革试点方案的成效不仅仅是结果的问题,它涉及到方案的制定、执行、落实(或享受)、监控、反馈等多个环节。只有从全过程、多维度来评估高考改革试点方案的成效,才能够发现改革过程各阶段与改革结果之间的因果关系,才能够科学而全面地考察高考改革试点方案的绩效和成败,从而为完善

高考改革试点方案提出更具针对性的改进、完善措施。

### 一、高考改革试点方案制定的评估

在高考改革方案分析中,方案制定是指方案出台的过程。高考改革试点方案的制定是最重要的一个环节,它决定着改革未来的成效,所以评估高考改革试点方案的成效不能抛开改革方案制定阶段,更不能忽视改革方案内容和改革方案文本。因此,我们可从两个方面评估高考改革试点方案制定环节:既评估高考改革试点方案制定程序,也评估高考改革试点方案文本。

#### 1. 评估高考改革试点方案制定程序

高考改革试点方案的制定是一个非常复杂的博弈过程,也是一个智慧凝结的过程,为此,应重点考察高考改革方案出台的时机、目标融合性、改革的准备情况、改革的民主参与情况以及方案制定的主体的层次性和利益独立性。

一是评估高考改革方案出台的时机。高考改革方案出台的时机非常重要,时机不到而出台改革方案,则达不到设计成效;时机到了不出台改革方案,同样会影响国家和社会的整体效益。方

本文系教育部人文社会科学研究规划基金“国家教育考试公平监测指标体系及保障机制研究”(项目批准号:14YJA880066)的阶段性研究成果。

余澄/华中师范大学考试研究院博士生,主要从事国家治理与考试制度、国际政治研究。(武汉 430079)

王后雄/华中师范大学考试研究院院长,教授,博士生导师,主要研究方向为教育考试与评价、课程与教学论。

# 国家教育考试公平标准及评价体系的构建

王世存 王后雄

【摘要】国家教育考试公平标准及评价体系主要包括选定评价项目、描述项目操作行为、划分水平刻度和分配项目权重系数等四大主体内容。其构建过程主要采用了直接经验法、专家咨询法、登记排序法和环比法。国家教育考试公平标准及评价体系共包括3个一级标目,27个二级标目,113个三级标目。公平标准及评价体系的构建为国家教育考试的公平、公正运行提供了参考和评判依据,对于推动我国考试制度改革,构建科学、公正的人才培养与选拔机制具有重要的现实意义。

【关键词】国家教育考试;公平标准;评价体系

【中图分类号】G424.74 【文献标识码】A 【文章编号】1674-1536(2012)05-0044-05

国家教育考试的公平、公正与安全不仅直接关系到考试本身的信誉与质量,关系到公共利益、社会秩序和公民的切身利益,也直接影响社会的稳定。党的十七大报告明确指出:“要始终把实现好、保护好、发展好最广大人民的根本利益作为党和国家一切工作的出发点和落脚点。”<sup>[1]</sup>因此,构建我国国家教育考试公平标准和评价体系既是解决国家教育考试公平的学理性当务之急,也是保障社会公平正义,促进社会和谐之现实需要。

## 一、国家教育考试公平标准及评价体系构建的基本步骤

国家教育考试公平标准是国家大型考试活动共同遵守的准则和依据,它以保证考试领域内的公正和质量为目的。国家教育考试公平标准及评价体系主要包括选定评价项目、描述项目操作行为、划分水平刻度和分配项目权重系数等四大主体内容,其中最为重要的是确定评价的标度和标准。因为确定标度和标准,既是国家教育考试公平标准及评价体系的重要理论问题,也是评价体系构建的科学方法问题,它在一定程度上影响着国家教育考试公平标准评价的科学程度。作为科学研究过程,构建标准体系必定历经确定评价目的、草拟评价标准、进行科学论证、广泛征求意见和组织试行修改等5大工作流程。从构建标准体系的业务技术角度来看,国家教

本文为教育部人文社会科学研究项目《国家教育考试公平标准及评价体系研究》(课题批准号:09YJA880049)阶段成果。王世存/华中师范大学教育学院博士生,信阳职业技术学院党委宣传部副部长,主要从事考试基本理论、课程与教学论研究。(武汉 430079)

王后雄/华中师范大学教师教育学院副院长,教授,博士生导师,主要从事考试基本理论、考试公平理论研究。

育考试公平标准及评价体系构建主要包括如下4个基本步骤。

### (一) 确定评价项目及其权重

首先,在明确国家教育考试公平标准及评价体系目标内涵、统一认识价值的基础上,确立国家教育考试公平标准及评价体系的总标目。然后运用考试评价理论,逐级分解总标目及次级标目,直至评价所需的末级标目。最后按照层级次序和层级内容次序对各级标目进行编码,并确立标目的权重系数。选择分解标目时,我们应重点考虑不同层级标目在国家教育考试公平标准及评价体系中的强度及频率,考虑影响标目评价质量的关键事件或关键表现,确保下级标目对上级标目具有良好的代表性和同质性。

### (二) 确定项目评价要求及其不同水平级别连续分布的标准刻度

所谓项目评价要求,就是要用直观性语言来描述末级标目的操作行为并体现其等次差异,其他层级标目则只需精炼的概念或短语作概括性表示。这一步骤的关键是要找出能够确切体现末级标目评价要求水平两极端语言变量,得到末级标目评价要求不同水平级别的连续分布体。如对“A11——设考目的明确、指导思想正确、依据充分”末级标目,我们选用“优”、“良”、“中”、“差”四种等次水平的评价要求,先找出两极端值的语言变量,然后可以得到从

# 不同利益群体对高考制度公平性认同度的调查与分析

王后雄 王世存

【摘要】高考制度公平性问题直接关系到教育领域的正义和公正,关系到社会的稳定和国家安全。本文通过不同利益群体对高考分省定额录取政策、高考分省命题考试政策、高校自主招生政策、普通高校保送生政策、高考加分政策、高考科目设置政策和高考命题取向政策等七个方面公平性认同度的调查与分析可以看出,要想真正实现高考制度的公平性,必须在高考理念方面进行变革、法制方面加强建设、机制方面着手创新、政策方面不断改进。

【关键词】利益群体;高考制度;公平性;认同度

【中图分类号】G632.474 【文献标识码】A 【文章编号】1674-1536(2011)09-0004-15

高考制度公平性是当前我国社会生活和教育领域的热点话题。它直接关系到教育领域的正义和公正,关系到社会的稳定乃至国家安全。<sup>[1]</sup>在当今中国,高考已成为“举国大考”。<sup>[2]</sup>高考制度公平性是教育公平中最直接、最重要的公平要素之一。美国学者戴维·伊斯顿(David Easton)认为:“政策是对全社会的价值做权威性的分配。”<sup>[3]</sup>高等学校入学招生考试政策实质上就是围绕高等教育入学机会的合理分配与调节所制定的行为规则。目前众多的高考不公平问题,基本都可以还原为高考制度问题。高考是一个涉及很多利益主体的系统工程,直接的利益主体有教师、家长和学生,隐藏其后的利益主体有社会、国家和民族。事实证明,我国的高考制度不仅起到选拔贤能的作用,同时在加快社会分层和人员流动、促进利益群体社会化、控制政治冲突和维系社会稳定方面都发挥着重要作用。因此,调查不同利益群体对高考制度公平性的认同度,从地域、城乡、民族、阶层等维度进行分析,不仅能为我国高考制度完善及高考改革决策提供科学的依据,同时可以为转型期我国高考制度的公平实施与运作提供制度和实践上的保障。

## 一、高考制度公平性的内涵和属性

公平是人类社会亘古不变的追求。而当考试成为一种社会的基因并且被当作诉求公平的主要乃至

唯一手段时,这个社会就成了一个考试社会。<sup>[4]</sup>高考制度公平性是公平的一种特殊类型,高考要发挥促进社会公平的功能,必须要以自身的平等为前提。当今社会,高考制度公平性已经成为教育现代化的价值取向,是世界各国教育立法的基本出发点。

### (一) 高考制度公平性的涵义

正确理解高考制度公平性的内涵需要从三个层面进行分析。

#### 1. 考试机会、考试权利和考试资源层面

高考制度公平首先是考试权利的公平,它承认并保证利益群体具有平等的考试权,这种权利不能受家庭背景、种族、性别以及资本占有状况等因素的影响和制约。<sup>[5]</sup>考试机会、考试权利和考试资源是三个不同的方面。考试机会表述的是一种概率,是社会利益群体考试的可能性。考试权利是一种人们的主观规定,是对社会利益群体考试可能性大小的设定。考试资源是一种客观条件,反映的是利益群体考试可能性的具体实现程度。

#### 2. 高考制度的起源和性质层面

高考不仅是利益群体的需要,也是社会的需要。高考制度公平性首先是一种价值判断和情感体验,它以利益群体的社会意识为前提,以社会的现实状况为对象,在一定的世界观指导下,对现实社会提供的教育机会、教育权利和教育资源分配情况作出主观评判。高考制度公平性也是一种现实状态,反映

本文为教育部人文社会科学研究项目《国家考试制度公平标准及评价体系研究》(课题批准号:09YJA880049)结题成果。

王后雄/华中师范大学考试研究院院长,教授,博士生导师,主要从事考试基本理论、课程与教学论及考试公平理论研究。(武汉 430079)

王世存/华中师范大学教育学院博士生,信阳职业技术学院党委宣传部副部长,主要从事考试基本理论、课程与教学论研究。

# 新课程标准下高考命题 若干技术难题刍论\*

王后雄

**【摘要】**新课标高考与选修模块对接的选做题的出现,选修模块与必修模块内容的选择,选修模块内容的多样性和难度,不同科目组分数之间的转换,选做题等值问题,新的考试科目设置、试题的选材及综合素质评价带来的不公平性等,都是新课标下高考命题亟待解决的技术难题。新课标高考必须有针对性地解决业已存在及新生成的问题,并尽可能避免产生更多的问题。

**【关键词】**课程标准 高考命题 技术难题

**【中图分类号】** G522 **【文献标识码】** A **【文章编号】** 1673-1654(2007)02-0031-11

教育部2007年新课标高考大纲已正式出台,与现行的高考大纲相比,新课标课程的实施决定了新高考的新特点:让学生全面而有个性地发展是这一次高中新课程改革的亮点,而新的大纲在维护统一考试的前提下,尽可能反映新课标教学中倡导的这种多样性和选择性,体现学生的个性化发展。教育部1999年“关于进一步深化普通高等学校招生考试制度改革的意见”中指出:高考内容的改革是高考改革的重点和难点,要花大力、长时间深入、细致地进行这项改革。作为高考改革的本质方向——高考内容的改革成为社会关注的焦点。本文力求从理论与实践方面探讨这一难点问题。

**作者简介** 王后雄,教授、主任,华中师范大学考试研究院。湖北武汉,430079。

\* 全国教育考试“十一五”规划课题“新课程下高考质量评价标准体系及其实施研究”阶段成果(课题编号:2006JKS1051)。

● 专家论考

## 高中学业水平考试制度:重构与审思

王后雄

根据《国务院关于深化考试招生制度改革的实施意见》,湖北省于2016年9月29日向社会公布了《湖北省普通高中学业水平考试实施办法》,该办法是湖北省“新高考”改革的重要配套方案。已公布的方案可概括为基于学业水平考试基础上的“统一高考科目(3)+选考科目(3)”科目模式,实行将统一高考、高中学业水平考试和综合素质评价纳入高校招生评价体系的“三位一体”新模式。一方面,坚持了统一高考的主体地位,另一方面通过分类考试、多元评价,丰富了评价立体和评价形式,倒逼高中学校总结和提炼发展学生核心素养的有效途径、方法和策略,从而提升自身开展“素养为本”课堂教学的能力,培养学生的立体意识、求异思维、综合素质、实践能力和创新精神。

1. 高中学业水平考试制度的审思。新的高中学业水平考试制度的设计是对传统学考制度的扬弃,在充分研究学考在高校招生录取中“硬”挂钩与“软”挂钩利弊前提下,充分表征学考和高考的关系用“和而不同”中的“和”来表示,同时基于高中教育对形成和发展国家学习能力的贡献而改进和完善学业水平考试制度。只有建立所认定的价值能独立、直接兑现的学业水平考试制度,才能从根本上确保高中学业水平考试制度以制度的形式存在,并发挥其制度的功效。

2. 高中学业水平考试制度的重构。高中学业水平考试制度的本质是国家用国家的学业水平兑现学生的学业水平,使学生学业水平成为国家学业水平存在的自然标志,学生凭借自己的学业水平得以自我实现。为兑现学生的学业水平必须建立起强制、权威、普遍的标准,使其具有全国统一的考试标准、规范的测评体系及严密可信的运作体系,成为一种独立的、具有高效度和高效率的、所认定的价值能直接兑现的权威性考试。

3. 跨越文理分科的两极思维模式。实行高中学业水平考试,旨在跨越文理学科之间的鸿沟,对全体中学生提出国家规定的所有学习科目的学业质量标准。这项改革将对我国21世纪人才的知识结构和思维模式产生深远的影响。可作为构

建国民素质的战略重点,在国民素质框架下设计学业水平考试体系,以期将其功能超越高中教育评价体系,而扩展到提升国民整体素质及国家人才战略层面上。

4. 高中学业水平考试的价值取向。学业水平考试内涵“扩张”至全科考试、全能考试、全体考试,从根本上保障学生发展的知识结构体系。选考科目的设计和一年多考机制的形成给学生考试内容和考试机会的选择权。将现行文理分科的套餐式科目设置改为必考科目基础上跨文理自主选科,从原来的2种(综合)科目设置选1扩大为20选1(上海、湖北等)、35选1(浙江),扩大了学生考试科目的选择权。部分科目增加考试次数,由学生自主选择具体考次,有利于学生在竞技状态较佳时参加考试,体现了尊重教育的个体需求。在测试和评价过程中给予学生主动发挥的空间,使学生从利益相关者进一步成为评价的主体。

5. 命题应以学科核心素养为宗旨。高中学业水平考试的主要目的是评价学生学科的核心素养的发展状况和学业质量的达成程度,这是区别于传统考试的关键特征。学业水平考试命题必须坚持以学科核心素养为导向,准确把握“素养”、“问题”、“情境”和“知识”四个要素在命题中的定位及相互联系。应依据学业水平考试中合格考试和等级考试的性质、特点及考试结果的应用价值,结合本地基础教育整体发展水平实际,科学、合理地确定试题的平均难度,高度重视试题难度的控制。

高中学业水平考试制度的建立,为从制度上将公平观念从形式公平升华到实质公平创造了条件。同时,我国反腐倡廉的正义清风和新媒体的高度发达为多元综合评价的信息透明提供强有力的监督措施,为保障规范和公平提供了制度和技术保证,为高考改革突破雷区、越过深水区奠定了坚实基础。

(本文系教育部人文社会科学研究规划基金“国家教育考试公平监测指标体系及保障机制研究”阶段性成果,项目批准号:14YJA880066)。

王后雄:华中师范大学考试研究院院长、教授、博士生导师,主要从事教育考试研究。

## ●专家论考

## 高考命题如何承载公平使命

王后雄

高考是我国高校科学选拔人才的大规模选拔性考试,因此,高考对公平性原则进行操作的\*\*首要任务,就是确定科学合理的选拔标准。命题对于高考公平影响最深刻、最广泛,而且一般情况下是隐性的、深层的,不容易察觉。高考命题对于高考是根本性的,对于高考公平也是关键性的。高考命题公平不是一个孤立的概念,而是在命题的各个环节中的公平性都必须考虑之中。高考命题不公平,高考实施乃至高考不可能公平。

从高考命题的确定过程来看,“考能力”成为我国当代选拔人才标准的新理念,这也是现代发达国家和地区高考改革的总体趋势。然而,高考过强的能力要求(如探究性学习、综合能力、创新实践等)与文化资本有关。文化资本通常很难在正式的学校教育中取得,大多通过父母优越的社会经济教育背景所创造的家庭教育\*\*与资源来形成。这些能力充家庭再生产的工具,使得考生资本存量的家庭差距通过高考的不平等代代相传。这表明以能力为导向的高考命题对于不发达地区及来自社会劣势阶层考生处境不利,一定程度上违背了考试公平的基本精神。高考命题要从知识与能力的关系上考量高考公平问题。

从高考试题内容的确定来看,高考对课程内容及其目标指向要求的\*\*畸轻畸重,即使考试区分度和信度很高,实际也包含着不公平。高考对科目的偏重是对相应科目具有学习优势和课程选择的学生的偏颇。部分省市高考科目方案加重了语文、英语等学科的比重,把通用技术和信息技术、艺术、综合实践等纳入高考的测试范围,势必对经济不发达地区和农村考生又是一种不公平。近几年,学者关注高考试题选材存在着一定程度的城市偏向,语文、英语、文综乃\*\*至理综试题情景材料较多反映城市生活,排斥乡土的文化与符号,从而对农村考生不利。为此,试题应避免出现城乡、性别及文化背景\*\*的歧视内容,如果试题中涉及不同群体的情景时,应该在试卷中体现出各种群体在测试中出现频次的平衡,综合考虑群体的多样性、差异性,将命题中的不公最小化。

从高考题型的选择来看,题型与考试的效度和信度有相当强的关系,题型的选择直接关系到命题的公平。每种题型都有其优劣,任何一

种题型都不能独立完成所有的教育测量目标。选择型试题有一定的猜测概率,信度较差,但内容样本较大,评分误差小;建构型试题能有效减少猜测的概率,信度较高,但内容样本较小,评分误差较大,评分信度较低。考生的认知特点不仅影响课程的学习,而且也影响他们对某一类题型的反应,从而造成与概念无关的系统性偏差。研究表明,场独立型的考生更擅长分散型试题(如英语听力、单项填空、完形填空、阅读理解等),而场依赖型的考生更擅长综合型试题(如英语书面表达)和口语考试。就英语学科来讲,高考英语题型对场独立型考生相对有利(分值\*\*达125分),对场依赖型考生的发挥空间形成偏见和\*\*不公平。因此,应从题型对信度的影响,对学生认知风格的影响,科学设计好各类题型及其比例,避免题型配置不当造成对公平及信度的影响。

从高考试卷难度的设计来看,高考试卷难度、学科之间难度匹配、选做题难度等设计不好,对高考公平也会产生显著影响。在进行试卷难度设计时,首先要平衡不同学科之间的难度关系。如果语文试卷偏易,数学试卷偏难,显然对于数学成绩较好的考生不利。此外,要设计好理科综合、文科综合试卷内部的难度结构。如果理科综合试卷中物理试题偏难,生物试题偏易,则对物理学科占优势的\*\*考生不利。对于新课程高考的选做题而言,为了保证高考的公平性,必须尽量保证不同模块的选做题的难度相近。这一技术难题在过去高考命题中不存在,是新课程高考多样性和选择性形成的难题,必须运用“等值”技术解决考试结果的公平性问题。从试题长度、知识点数、能力层次、难度系数、解答题时间、答题字数等可比性技术要素相近,用测验等值的技术保证编制的选做题是平行题。

高考命题的背后,承载着社会各方面的利益期待和价值诉求——文明的期望、教育的责任、公众的功利、学生的前途、社会的需要,命题公平的本质在于实现个体人的价值,消除城乡偏见和阶层偏见,平等对待各类考生。要实现命题公平必须为命题公平的贯彻提供制度保障,公平作为命题工作者实践创作的首要目标,必须纳入审查和监督范围。建立命题首要目标审查制度,着力提高命题者的公平素养,才能让高考命题承载其应有的公平使命。

王后雄:华中师范大学教师教育学院副院长、教授、博士生导师,主要从事教师教育及教育考试研究。

## ●专家论考

## 健全高中学生综合素质评价制度改革

王后雄

《国家中长期教育改革和发展规划纲要(2010—2020年)》中指出,普通高中要建立科学的教育质量评价体系,全面实施高中学业水平考试和综合素质评价。2014年9月,国务院印发《关于深化考试招生制度改革的实施意见》(以下简称《实施意见》),两大文件的出台,标志着新一轮考试招生改革的全面启动。这是中国教育领域一项具有基础性、全局性意义的重大制度改革。在高考综合改革方案相继出台,在维护“考试选拔”这一基本取向的基础上,确立了“分类考试、综合评价、多元录取”的改革模式,其中“综合素质评价”是本次招考改革的一大亮点,要解决的主要问题是教育制度性问题导致的“唯分数论”“一考定终身”,影响学生全面而有个性化的发展。其目的是探索招生与高中学习相关联的办法,增加学生选择的\*\*灵活性,分散学生的考试压力,发挥考试招生在引导素质教育、促进学生全面发展中的重要作用。探索基于统一高考和高中学业水平考试成绩,参考综合素质评价的多元录取机制。这不仅会在促进教育公平、科学选才方面发挥重要作用,而且必将引领和推动教育领域的综合改革向纵深发展,其意义重大而深远。

随着基础教育课程改革的不断深入,人们更\*\*加深刻地认识到教育评价已经不能仅仅停留在学生学业成绩的评价上面,还必须更加注重学生综合素质的培养与评价。人才的选拔必须从只看“冷冰冰的分”到关注“活生生的人”;评价的主体不再局限于教师一人,学生的自评和互评、家长和社会的评价都应被\*\*全面纳入;评价的内容不再仅局限于知识获得,而应更加多元化,道德品质、公民素养、学科能力、交流与合作品质、健康水平、审美表现等维度都应该成为评价的内容;评价的方式不再局限于单一方式,形成性评价和终结性评价的结合,定量评价和定性评价的结合,纸笔测验与活动表现评价、档案袋评价的结合,都应成为评价学生综合素质的重要方式。科学地检验教育活动的效果,促进学生在综合素质方面的全面发展,进一步扭转“唯分数论”的评价观是当前健全高中学生综合素质评价制度改革的主要动力与方向。也是考试招生科学性\*\*与公平性的实现内在统一的价值诉求。

《实施意见》提出了推进学生综合素质评价的四个基本原则,即以社会主义核心价值观为\*\*中心的方向性原则,注重“以人为本”的\*\*指导性原则,以事实为依据并真实记录学生成长过程的客观性原则和严格规范、有效监督的公正性原则。同时,《实施意见》还将学生综合素质评价内容分为思想品德、学业水平、身心健康、艺术素养和\*\*社会实践五个方面,并明确了每个方面的考察重

点。针对之前的综合素质评价制度“操作性不强,难以把握”的不足,《实施意见》进一步强调了综合素质评价需要注重考察学生的行为表现,特别关注于通过学生在有关活动中的具体表现来反映其全面发展情况和个性特长。

然而,教育的改革不可能一蹴而就,且往往都是一个螺旋递进的过程。虽然评价学生综合素质的观点早在本世纪初就已被提出,然而,在实践中都会面临一些问题,如教育现状看,认识不到位、操作性不强、信度不高、监管力不强,结果不真实、公信力差等问题仍然屡见不鲜。作为制度的设计者和实施者,有效地解决这些问题,既需要科学认识评价的客观规律以提升评价改革的理论理性,也需要探索如何评价才能更好地满足人的发展需求以提升评价改革的实践理论。启动高中学生综合素质评价,主要是要完善评价规则、改革监管机制。通过加强信息公开、加强制度保障、加大违规查处力度,才能保障评价公平公正。

健全高中学生综合素质评价制度改革,关键是要杜绝可能发生的“人情评价”“关系评价”“\*\*金钱评价”等现象,避免人民群众普遍担心的“评价不公”问题的出现,确保社会各界在高校招生入学中的平等地位。为了进一步深化改革,要不断发现新的问题,解决实施中的新问题。首先,要研究如何将高中学生成长过程中的突出表现作为考察的重点,如何通过参与相关活动情况及其成果来考察学生的综合素质状况,从而使评价内容可考察、可比较、可分析。其次,要研究\*\*写实记录、公示审核等评价机制的建构和运行,从而\*\*保证如实记录学生成长过程中的具体活动,并以\*\*事实材料为佐证,做到有据可查。再次,要研究如何建立健全材料公示制度、抽查制度、申诉及\*\*复议制度、诚信责任追究制度等,如何对弄虚作假\*\*按国家有关规定给予严肃处理,从而确保综合素质评价材料真实可靠。

实施高中学生综合素质评价制度,凸显了更\*\*高水平的科学性\*\*与公平性的交融,与教育改革\*\*方向及人的个性化发展相适应,在最大限度上持续释放教育评价内在的魅力。可以预见,即将付诸\*\*实施的高中学生综合素质评价制度必将给中国社会及中国教育带来深刻的变化与深远的影响。对于高中学校而言,在面临新挑战的同时也迎来了\*\*自身改革创新、转型升级、定位特色的契机,为学生的个性化发展创设更为丰富多元高质量的空间、平台、机会、资源,为学生提供更\*\*为专业的职业生活教育指导,这才是对高考综合改革方案的最好回应。

[本文系教育部人文社会科学研究规划基金(项目批准号:14YJA880066)的研究成果之一。]

王后雄:华中师范大学考试研究院院长、教授、博士生导师。

●专家论考

## 精准落实立德树人考查目标 彰显高考评价体系实践自信

王后雄

2017年的高考注定引人关注,因恰逢恢复高考40年,又是上海、浙江全面实施高考综合改革元年,我国高考已步入由传统高考制度向新高考改革的关键性过渡期。而2017年高考命题改革显然成为人们识别高考改革的风向标,社会将高考聚焦于能否满足社会对人才培养的全方位、多角度的要求,真正办成让人民满意的高考,能否为国家层面深化考试招生制度改革提供借鉴推广范本的高考命题改革之路。

**第一,目标与功能:高考命题的精准定位。**2017年高考各学科命题延续了多年来业已形成的优良风格,也展示出课程改革与高考改革互为关联的新的亮点。《国务院关于深化考试招生制度改革的实施意见》颁布后,历经三年的不懈探索,今年进入了全面推进阶段。2017年高考命题全面贯彻党的教育方针,试题精准落实立德树人的考查目标,以服务高校人才选拔、导向中学教学为命题出发点,在继承中华优秀传统文化传统中,彰显文化自信。命题较好地体现了继承与发展、稳定与创新的关系,以提升考试的科学性和公平性为目标,依据高校人才选拔要求和国家课程标准,优化考试内容,体现素养导向,增强基础性、综合性、应用性和创新性,适应经济社会发展对多样化高素质人才的需要。考试命题的价值取向随着我国教育改革和高考改革的大环境而发生演变,这种演变直接影响了考试目标、考试内容、考试形式和命题思路的发展方向。纵观2017年高考,回应了教师、学生和社会关切,保证了高考平稳顺利过渡,发挥了精确的育人选拔功能和积极的导向作用。

**第二,能力与素养:高考命题的质量引擎。**依据2017年高考学科考试大纲的修订与调整改革要求,为完善考核目标,结合了学科特点和核心素养的内涵,实现了“关键能力”与“学科素养”的全覆盖。语文将文学类与实用类文本均设为阅读必做题,实现对不同思维方式、素养构成有别的考生全方位考查,信息筛选、逻辑分析、审美鉴赏及语言运用等能力的全面覆盖。数学以基础知识、基本技能、基本思想方法为支撑点,加强对理性思维、通性通法的考查,突出对创新应用能力的考查。英语试题坚持以考查综合语言运用能力为目标,强调学科的基础性、综合性、应用性和创新性,弘扬中华优秀传统文化,引导学生加强对人类命运共同体的关注和科学精神的养成,突出命题

的育人功能。理科综合体现学科特色,试题素材丰富,重视对思维能力、实验能力和科学素养的考查,试题联系生产、生活实际,注重应用能力、创新能力的考查。文科综合试题厚重大气又不足活力,时代感鲜明,突出社会主义核心价值观、中华优秀传统文化考查,反映国家建设发展成就,引导考生增强“四个自信”,突出学科主干知识、学科素养和关键能力的考查,发挥引导中学教学的积极作用。

**第三,公平与科学:高考命题的价值追求。**2017年高考命题秉持促进公平、科学选才的价值取向,不断改革完善。高考各学科逐步完善了知识、能力、素养的考查目标及层次标准,使之考查学科化、具体化。同时优化测量目标及题型功能,调整了必考与选考的学科内容,完善了试卷结构的设计,在命题环节较好地体现了知识、能力、素养的辩证关系,提升了高考选拔的有效性。考虑我国国情和教学实际,高考命题并没有完全照搬国外标准化测验模式,转而研发信息迁移题、工艺流程题、材料解析题、材料作文等新题型,合理使用符合学科特点的题型评价相关能力、素养及其水平,力求避免猜题、押题、宿构,有利于真实评价考生能力水平,体现了我国高考评价体系的实践自信。按照教育测量学理论,大规模常模参照性测验的试卷难度在0.50左右时对考生总体的区分度较大,综合考虑各种因素,随着我国高校招生录取率的提高,现行阶段试卷总体难度以控制在0.55~0.65之间较为适宜,2017年多数学科难度控制是合理的。试题选材是否涵盖城乡、民族、区域、性别的差异是命题公平应关注的要素,今年高考试卷较好地避免了城市中心取向、性别倾向等公平问题。客观来说,2017年高考命题在把握试题不超纲、试题难度不能高、区分度不能低、试题选材不能有失公平、重点难点不回避、试题保持原创的前提下,有效地控制了考试误差,保证了高考公平、科学选才。

总之,2017年高考命题坚持整体稳定、推进改革创新、顺应课程标准修订的趋势,较好地体现“必备知识、关键能力、学科素养、核心价值”的层次与关系。然而,高考命题改革如何全面充分地考量理念冲动和现实约束的关系,遵从在公平与科学上统筹的现实理性,兼顾个体价值与社会需求的价值取向,体现高考“立德树人、服务选拔、导向教学”的根本立场,还需要展开更加深入的研究。

王后雄:华中师范大学考试研究院院长、教授、博士生导师,主要从事课程与考试改革研究。

## 基于化学核心素养的高中学业水平考试命题策略

王后雄

(华中师范大学考试研究院,武汉 430079)

**摘要:**化学学科核心素养构成要素之间具有内在的本质联系,根据高中学业水平考试的特征,化学学业水平考试命题必须以化学核心素养为导向,准确把握“素养”、“情境”、“问题”和“知识”四个要素在命题中的定位及相互联系,围绕必备知识、关键能力、学科素养、核心价值的考查内容,建构以化学学科核心素养为导向的命题框架。

**关键词:**高中化学;学科核心素养;学业水平考试;命题策略

**中图分类号:**G633.8 **文献标志码:**A **文章编号:**1000-0186(2018)04-0115-09

2014年9月,国务院颁布《关于深化考试招生制度改革的实施意见》(以下简称《意见》),就深化考试招生制度改革作出系统设计和全面部署。教育部随后出台《关于普通高中学业水平考试的实施意见》等4个配套性文件,正式拉开了新一轮考试招生改革的大幕。《意见》明确指出,不仅课程标准需要基于核心素养来设计和实施,高中学业水平考试的学业质量标准 and 考试标准同样需要围绕学科核心素养来进行设计,同时将高中学业水平考试中的部分科目直接作为高考科目计算总分,这无疑是一种崭新的制度设计,将是这一轮考试招生制度改革成败的一个关键。

**基金项目:**本文系教育部人文社会科学研究规划基金“国家教育考试公平监测指标体系及保障机制研究”(项目批准号:14YJA880066)、华中师范大学重点项目“学科教学论教师教学实践共同体协同模式与实践研究”、华中师范大学教师教育重大科研课题创新示范基地培育项目“学科教学论教师实践共同体持续发展的有效机制研究”的阶段性研究成果。

**收稿日期:**2018-01-17

**作者简介:**王后雄,1962年生,男,湖北黄冈人,华中师范大学考试研究院院长、教授,博士生导师,主要从事化学课程与教学论、考试与评价研究。

### 一、化学学科核心素养概念的界定及关联要素分析

#### (一)化学学科核心素养概念的界定

化学学科核心素养是指学生在接受化学教育过程中逐步形成的适应个人终身发展和社会发展所需要的正确价值观念、必备品格和关键能力,<sup>[1]</sup>是学生通过学习化学内化的、带有学科特性的品质。其中,“正确价值观念”指向学科育人的价值,“必备品格”指向学生内在的修养,“关键能力”指向学生外发的力量。高中化学学科核心素养是学生发展核心素养的重要组成部分,是高中学生综合素质体的具体体现,是学生科学素养的关键成分,反映了社会主义核心价值观下化学学科育人的基本要求,全面展现了化学课程学习对学生未来发展的重要价值。

王后雄教师教育系列教材

# 化学教育 测量与评价

Huaxue Jiaoyu  
Celiang Yu Pingjia

王后雄 李佳 /主编

王后雄教师教育系列教材

# 教育考试的 理论与方法

Jiaoyu Kaoshi De  
Lilun Yu Fangfa

王后雄 /主编

王后雄教师教育系列教材

# 新理念化学 教学诊断学

Xinlianian Huaxue  
Jiaoxue Zhenduanxue

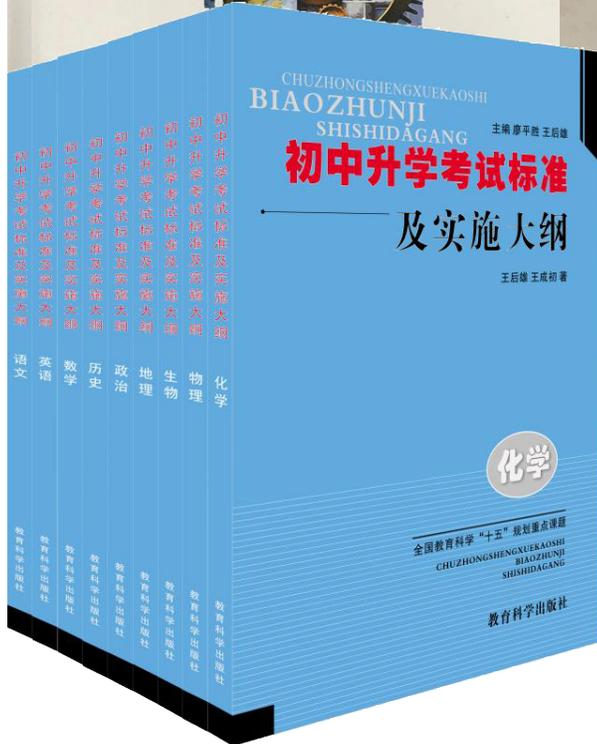
王后雄 王世存 /编著

卓越教师教育精品丛书

# 中学化学 奥林匹克竞赛研究

王后雄 主编

科学出版社



感谢聆听！