

化学高考 标准化知识

广东省化学高考标准化研究室编



广东高等教育出版社

化学高考标准化知识

广东省化学高考标准化研究室编

广东高等教育出版社

化学高考标准化知识

广东省化学高考标准化研究室编

*

广东高等教育出版社出版

广东省新华书店发行

广发印务有限公司印刷

787×1092毫米32开本 4印张 81.8千字

1986年12月第1版 1986年12月第1次印刷

印数1-100,000册

统一书号7343·33 定价0.80元

前言

随着我国高校招生考试改革的发展，广东省进行标准化考试试验的科目有所增加，除已进行的标准化考试的英语、数学和物理科外，化学科也将在1987年列为标准化考试试验科目。

为了搞好这项试验工作，成立了由化学和教育测量方面的专家与教师组成的广东省化学高考标准化研究室，进行了大量的研究和准备工作。初步拟定了化学高考标准化考试大纲、命题双向细目表，进行了标准化考试预测，并正在建立标准化考试题库。本书就是在该室研究工作的基础上编写而成的。它将给读者介绍化学高考标准化的基本知识和内容，从而为中学化学教师 and 准备参加1987年化学高考标准化考试的我省考生，提供必要的指导。本书可供普通中学化学教师和学生阅读，也可供其他中等专业学校师生参考。

全书由两部分组成，第一部分介绍化学标准化考试的基本知识；第二部分则为练习题，含一份曾作为预测使用的标准化考试试卷，可供读者作练习参考用。

参与本书编写的同志有梁远明、曾桂兴、李学尧、刘仁美、杨世柱、凌燕湘、曾灼先、张多霞、徐桂荣、梁尧，由谢晋显负责绘制插图，全书由江琳才负责统编。此外，还有许多高校和中学有经验的教师和同志，参加过本书的某些部分的讨论和修改工作，提出了许多宝贵意见，在此，谨向他们表示热诚的谢意。

由于标准化考试还处于试验阶段，我们对化学标准化考试的认识还十分肤浅，而且编写时间仓促，因而书中肯定会有不足和错误之处，恳请读者批评指正。

编 者

1986年8月于广州

目 录

第一部分

- 一 教育测量与标准化考试..... (1)
- 二 传统考试的弊端..... (5)
- 三 评价考试质量的若干指标..... (7)
- 四 化学高考标准化的标准..... (14)
- 五 题型与试卷设计..... (15)
- 六 考试目标..... (18)
- 七 双向细目表..... (21)

第二部分

- 一 选择题练习..... (25)
- 二 综合练习题..... (87)
- 三 参考答案..... (98)
- 四 一份化学高考标准化考题试卷..... (104)

第一部分

一. 教育测量与标准化考试

标准化考试是属于教育测量学的研究范畴。教育测量学是一门对教育对象及其成效进行科学测定和评价的教育学科。它包括对人的知识、技能、品格、兴趣、爱好、学术倾向等方面的研究与测定。学习成就方面的标准化考试仅是其中的一个分支。

教育测量不同于我们通常所说的物质测量，后者指的是直接性测量（如物理、化学上对有关物质的测定），前者则为间接性测量。例如，我们要了解一群学生所掌握的化学知识和技能，除进行个别的操作性（如化学实验）观测（这在大规模考试中是无法做到的）之外，往往是通过一份编造好的化学试卷对考生所具备的化学知识和技能进行测定和评价，这就是教育上的间接测量。这种间接测量存在着三个基本问题：(1)这种测量有无一定的标准？这是测量标准问题；(2)是否测量到我们想要测量的东西？这是测量效度问题；(3)测量结果准不准？这是测量的可靠性或称信度问题。这三个问题是教育测量上至关重要的问题。

从教育测量的观点来看，任何一个考试都是测量考生学习成就和技能的一把“尺子”。然而这把“尺子”造得怎么样，是否合乎某些既定的标准，是否符合一定的性能指标，这些都是这把“尺子”的基本问题。所谓标准化考试，是指要造一把比较标准或称比较规范性的“尺子”，用这把“尺子”测量学生时，能符合一定的指标(如效度、信度、难度、区分度及测量误差等)要求，测量结果能够比较真实地反映考生的实际水平。从理论上说，标准化考试是对考试制定出客观而规范性的标准，从命题、实考、阅卷、评分、计分等各个环节都努力减少或控制各种误差，从而测出考生比较真实成绩的系统过程。

标准化考试大概有这么几个要求：

1. 要有考试大纲或考试指导书，用以规定考试的目的、内容、要求、方式、题型及评分与计分的方法等。考试大纲由专家和有经验的教师根据教学大纲共同制定，并经教育主管部门批准，及早公布于众。这样做的目的是使广大考生、学校都明确考试的要求与方向，都向同一目标努力，因而有利于减少随机误差，测出考生的实际水平；其次，由于考试大纲是教学大纲的具体化，因而使教师对教学大纲具有真实感，有助于教师努力落实教学大纲，因此考试大纲对平时教学能起参照和促进作用，有利于从整体上提高教学质量。

2. 命题要标准化。命题的依据是根据考试大纲而制定出来的“命题双向细目表”(详见第七节)。下表就是我们考虑的化学高考命题双向细目表：

表1-1 :化学高考命题双向细目表

考试目标 考试内容	识记	理解	应用	分析 综合	探究	总分
基本概念 基础理论	1	13	9	7	2	32
元素化合物	3	5	6	5	2	21
有机化合物	1	5	3	4	2	15
化学计算		3	4	8		15
化学实验	1	6	2	6	2	17
总分	6	32	24	30	8	100

这是一种考试目标与考试内容的双向列联表。表中各细格的数字具体地规定了各考试目标与考试内容之间的比例。如考试目标在“识记”方面占6%，其中“基本概念”占1%，“元素化合物”占3%，“有机化合物”占1%，“化学实验”占1%；

在考试内容的“化学实验”方面占17%，其中“识记”占1%，“理解”占6%，“应用”占2%，“分析与综合”占6%，“探究”占2%；余类推。命题者要据此来命题，命题后还要据此来评估试题的覆盖面（即后面讲到的内容效度）是否符合既定的要求。

3. 试题要经过预测和调试。●通过预测可以发现考生对试题的各种可能回答，以便制定更客观的评分标准和计分措施；同时，可以检查命题质量，删去或修改试题中干扰性太差的选项；还要计算各题的难度和区分度，进而评估整卷的信度与效度。经过预测后对题目进行筛选，保留合乎质量要求的题目，淘汰或修改不合质量要求的题目（如删去或修改区分度太低的题目）；在编制试卷时，对选用的题目要按难

度从易而难地排列，使整份试卷难度适中、区分性能好，使考生实测成绩保持正态分布。

4. 考试的组织实施过程要统一而规范。试卷的印制与输送要严格保密；实考的时间、环境要统一；监考员在试室只能讲规定要讲的话；要严格按事先规定的评卷步骤和评分要求来组织阅卷评分；尽量使用电子计算机来阅卷评分等。这些措施的目的还是为了减少或控制测量误差，使考试分数有良好的可靠性和可比性。

5. 要提供为解释考试分数之用的常模。所谓常模是指测量的常规量数，如考试分数的平均数（代表考试分数集中位置或水平的量数）、标准差（代表考试分数分散状态或称波动状态的量数）和各种相对地位量数（如百分等级）等。标准化考试的平均数和标准差是稳定的、相对保持不变的。有了常模，可以使考生了解自己的成绩在考试总体中处于什么水平，使考生成绩之间、学校考试成绩之间、地区考试成绩之间可以进行比较。此外，对于涉及多科考试成绩的综合，要利用各科的常模资料把原始分数转换成标准分数，使考生各科考试分数具有可加性，使考生总分具有可比性。

一般而言，标准化考试有如下特点：考试目标明确，考试内容全面，测试水平稳定；建立题库并按一定的工作程序和一定的质量指标制定试卷；试题取样范围大，题量大、覆盖面宽，考试的效度和信度都较好；试题难度适中，区分不同程度考生的能力较强；试题答法尽量简单、明确，使评分客观、准确；从命题到阅卷评分等若干环节都努力减少无关因素的影响，使考生得分可靠。总之，标准化考试，从命题到考场的组织安排，以至答卷的评阅、考试成绩的分析与评定，都

非常讲究科学性。标准化考试的本质是建立各种测量标准、减少或控制各种测量误差。

可见,标准化考试是一种较科学或称较客观的考试方法。从上述标准化考试的性质来看,化学科完全可以按照标准化考试的程序来实施。调集教育测量专家和化学科专家及有经验的教师,共同研究化学科的知识与能力的关系,明确考试目标(包括知识与能力的层次及其关系)和测量标准,探讨各种题型的结合,探索具有我国特色的化学高考标准化的路子,这是当前化学科标准化考试的最迫切任务。

二、传统考试的弊端

传统考试亦称旧法考试,是一种主观性和经验性的考试方法。这种方法通常是由主讲教师根据教学内容及要求,出三、五道,多则十来道试题,然后教师根据自拟的评分标准和学生的书面答案给定一个分数。这种方法在一般的教学班级中使用是有实用意义的,所以,这种传统考试方法在一般教学班级中仍广泛使用。但是,如果把这种方法推广到较大范围,或延伸到竞争性或筛选性较强的考试中,其弊端就会异常突出了。

经验是宝贵的,但经验还不是科学。回顾我国传统性高考,可以发现其中确实存在不少缺陷,如:试卷的制定没有明确的标准;阅卷方法落后、时间长;将卷面分数直接加入总分不够科学等。在国外有不少的教育学者进行过专门的研究。例如,著名的教育家斯达奇(starch)曾做过这样的实验:他把同一份英文答卷,分请一百四十二位本科毕业的中学教师评分,结果对这份答卷的给分有三十五种,从50分至98分

高低不等；他又把同一份几何答卷分请一百一十六位本科毕业的教师评分，结果得分六十多种，最低为28分，最高为92分。为了研究评分误差情况，笔者在1984年对我国高考理科（七科）的评分作过类似的调查和实验。实验的方法是：以学科为单位，随机地抽取五名考生的答卷，分别复印十份，分请十位高考阅卷者根据既定的统一评分标准给分，从中考察不同评分者对同一份答卷给分的差异情况。实验结果表明，各科均出现不同程度的差异，即所谓评分误差。误差较大的是语文科（五份考卷平均误差11.5分）和政治科（平均误差为6.5分），其他理科的平均误差都不低于5分；误差较小的倒是当年全都使用客观题（指答案简单或唯一的题型）的英语科，平均误差才1.2分。为了进一步了解作文题的评分误差情况，笔者把其中一篇十人评分误差已为10分的作文，再复印二十份，分请另外二十位评分者给分，结果对这篇作文竟有十五种分数，评分误差达20分。还可以估计，随着阅卷给分人数的增加，对这篇作文的评分误差还会加大。传统考试方法在高考中的弊端可见一斑。

总的说来，传统考试方法弊端主要是：经验命题，主观定评分标准；题量少、覆盖面窄；不注重对考试质量（指效度、信度、难度和区分度等）作数量化的分析；评分误差大等。

这些弊端极不利于高校选拔人才，也不利于提高中学教学质量，必须尽快克服。克服的办法就是设计出一套客观的考试方法，实行标准化考试。

标准化考试与传统性的教师自编考试的对比，可归纳为下表所示：

表1—2: 标准化考试与传统性考试的比较

	标准化考试	传统性考试
测量的学习内容 与结果	测量一般的学习结果, 内容适合于大多数学校; 属一般技能和理解力的测验, 并不反映当地教材具体而独特的重点。	良好地适应当地课程具体内容而独特的内容与结果; 能适应各种教学单位, 但有忽视测量复杂学习结果的倾向。
题目的质量	题目质量一般较高; 题目由专家编写, 并经过预测, 根据项目分析结果进行筛选。	题目质量常常是未知的, 一般比标准化考试低。
信度	信度一般在0.8~0.95之间, 常常在0.90以上。	信度通常是未知的。若仔细编造题目, 则信度可以提高。
实施与评分	实施步骤标准化, 提供特定的指导语。	可能统一步骤, 但通常是灵活的和不成文的。
分数的比较与解释	考试分数与常模比较, 另有考试手册帮助解释考试结果。	分数的解释与比较只限于本班或本校。

三. 评价考试质量的若干指标

教育测量学告诉我们, 评价考试质量不能仅凭主观评估, 而必须以数量指标来分析, 以实际效果来检验。评价一次考试的质量常用下列指标来分析和评价。

1. 效度 (validity) : 这是考试有效性或正确性的指标, 指考试的有效程度, 即是考试能够反映我们想要测量的学习结果或学习特质的程度。例如一把尺子, 用来测量人的身高

是很有效的；但若用来测量人的体重，那效果显然是很差的。又如化学高考，应是测量考生的化学知识和能力水平的，但如果我们把考题出得隐晦曲折，那首先就会变成对考生阅读理解（语文知识）的测量；或者在化学考试中含大量生物方面的题目，考生运用已有的生物化学知识就能得到高分，这种考试就会变成生物化学知识的测量。显然，这样考试的效度是不会好的。

效度可分为内容效度、目标相关效度和结构效度等。

内容效度就是通过编制如表1—1那样的“命题双向细目表”来体现，并通过编制“试题分类表”来检验。根据表1—1编制出相应的试题分类表的式样如下表1—3所示：

：据表1—1的试题分类表

题号	考试内容	识记	理解	应用	分析综合	探究
1	基本概念	✓				
2	有机化学	✓		✓		
3	基础理论	✓	✓			
4	化学计算				✓	
⋮	⋮					
m	基础理论		✓			✓
	合计	8	30	24	31	7
	覆盖率	133%	94%	100%	103%	88%

这种试题分类表主要靠学科专家根据既定的考试目标和考试内容，对试题逐一评估。例如第2题，化学学科专家认为是

考了“有机化学”的内容，在“考试目标”上他认为既考了“识记”，又考了“应用”（在他认为考了什么东西下面打“√”即是）；第1题则在“基本概念”方面只考了“识记”；余类推。这种试题分类表最好由若干个学科专家来评估（或计其平均数）。最后计算一下各个考试目标的覆盖率，即试题分类表覆盖命题细目表的情况。如在“识记”方面，命题细目表中要求占6%，试题分类结果占8%，则覆盖率为 $8\%/6\% = 1.33$ ；余类推。在化学高考中，要求覆盖率在0.9以上。显然，有了这种内容效度，我们就可以认为在考试（包括考试目标和考试内容）考到了我们想要考的那些东西。

目标相关效度是指考试结果与某些考试目标相关连的程度。每一考试都是与一定的目标相关连的，如高考是为了选拔中学毕业或具相当水平的优秀人材进入高校深造；自学考试是为了发现那些虽未经过正规院校学习（即通过自学）而具有同等学力的人才；考核干部是为了使用干部等，那么我们就可以从高校是否选到有用之才，从有关院校毕业水平，从使用干部中来检查这些考试的效果。例如，我们可以通过计算高考化学成绩（X）与大学化学专业一、二年级学习成绩（Y）之间的相关系数来体现化学高考的效度。这种效度就是目标相关效度。这就是说，如果化学高考很有效的話，那么在化学高考中成绩较高的学生应比成绩较低的学生在专业学习方面更有成就，即应呈现出显著的正相关。相关系数的计算公式为

$$r = \frac{N\sum XY - (\sum X)(\sum Y)}{\sqrt{[N\sum X^2 - (\sum X)^2][N\sum Y^2 - (\sum Y)^2]}} \quad (1)$$

结构效度是指考试能够测量到心理上某些结构或特质(如推理能力、综合能力、评价能力、实验能力等)的程度;亦即依据心理学上的结构来解释或分析考试分数所代表的意义。评价结构效度的主要方法是教育统计上的因素分析法,由于此法较为复杂,故在此略述。

2. 信度 (reliability):这是考试可靠性或稳定性的指标,指考生成绩的可信程度。考试结果应是知识与技能好的学生能得到高分,学习水平较低的学生得到低分;而且不管测多少次,这种有序性的结果都相对稳定,则称考试可靠或可信。信度值越高,表明考试结果越可靠;信度值偏低,表明考生得分的随机成份较大。决定一次考试信度的理想方法是,在相同的情景下用这个考试对同一批考生考两次以获得两组分数,进而比较这两组分数的差异情况:误差愈小,则考试信度愈高。但在事实上,我们几乎不可能在完全相同的情况下去实施这个考试。因此,要用其它方式来估计考试信度。较常用的方法有:

(1)折半法。在命题时就有意识地把每项要考查的内容与目标都分成等效的两半(如一半为奇数题,另一半为偶数题,这两半的考试内容、考试目标、难度大致相同);考试后计算每个考生在这两半试卷得分之间的相关系数 r_0 (用公式(1)),然后用下列的公式计算整个考试的信度:

$$r_u = \frac{2r_0}{1+r_0} \quad (2)$$

式中 r_u 为全卷信度, r_0 为两半卷之间的相关系数。

(2)库——李方法。如果考试题型全为选择题,则可用下面公式计算整个考试的信度:

$$r_{KR} = \frac{m}{m-1} \left(1 - \frac{\sum pq}{S_T^2} \right) \quad (3)$$

式中 m 为整个考试的题数, S_T^2 为每个考生得分的总方差(标准差的平方), p 为每题中答对的比例即

$$p = \frac{K}{n} \quad (4)$$

n 为考生总数, K 为某题答对的人数, $q=1-p$ 。

若考试题型不全为选择或全为论述题时, 则可用另一公式计算考试的信度:

$$r_a = \frac{m}{m-1} \left(1 - \frac{\sum S_i^2}{S_T^2} \right) \quad (5)$$

式中 S_i^2 为考生在第 i 题上得分的方差, 其它符号同式(3)。

3. **难度 (difficulty)**: 这是指试题的难易程度, 它是评价试题质量的重要指标之一。难度的计算公式为: 若题目为选择题, 则可用公式(4)计算; 若题目为非选择题时, 则用

$$P = \frac{\bar{X}}{W} \quad (6)$$

来计算每题的难度, 式中 $\bar{X}$ 为考生在某题上得分的平均数, W 为该题的满分值。

4. **区分度 (discrimination)**: 它用以评价试题对不同考生的知识与能力的鉴别程度。由于高考的重要功能之一是把不同程度的考生区别开来, 以利择优录取, 所以试题的区分度也是衡量试题质量的重要指标。

*该公式计得的结果实为题目的易度, 真正难度为 $q=1-p$, 但国际上常用该公式来表示题目的难度, 故本书亦采用这种公式。