

标准化考试的理论与实践

廖平胜

何雄智

编著

华中师范大学出版社

24.7

标准化考试的理论与实践

廖平胜 何雄智 编著

华中师范大学出版社

标准化考试的理论与实践

廖平胜 何雄智 编著

华中师范大学出版社出版

(武昌桂子山)

新华书店湖北发行所发行

孝感报社印刷厂印刷

开本787×1092 1/32 印张5.25 字数120,000

1986年12月第1版 1986年12月第1次印刷

印数00,001—15,000

统一书号7406·35 定价1.00元

绪 言

《标准化考试的理论与实践》，是一本阐述在我国社会经济及产业结构急剧变革的情况下考试如何改革的书，是一本论证在传统考试具有悠久历史的中国要不要实行标准化考试的书，是一本系统介绍如何设计和实施标准化考试的书。

广泛普及标准化考试的理论和实践知识，聚集各方智慧共探考试之新法，在合理继承我国传统考试科学成份、择善吸收国外考试先进经验的基础上，尽早创立适合我国社会发展需要的现代标准化考试体系，使考试更好地服务于祖国的建设大业，乃是本书的宗旨。

中国素有“考试故乡”的誉称，随着社会的演变，考试在我国走过了一条极为漫长的发展道路。它作为一种测度、甄别人的素质和智能个别差异的方法，一种评价、选拔、任用人才的手段，早已被普遍用于社会生活和社会生产的各个领域。我国的传统考试可谓人所共知。但什么是标准化考试？怎样设计和实施标准化考试？它与我国传统考试区别何在？为什么我国要尽快实行标准化考试？却是我国绝大多数人极为陌生而又渴望知晓的问题。

近些年来，随着我国经济的发展和科学技术的进步，社会经济结构、产业结构、就业结构相应变革，传统考试越来越不适应现代化建设对人才培养、选拔和任用的需要，急待除旧布新。因此，富有时代色彩的标准化考试，在我国日益引起人们的关注，社会各界为之呼吁、探究者甚多，有关标准化考试的资料、文章大量见诸于报刊，标准化测试手册、文

集，试题集争相印行；标准化考试的试验范围逐年扩大。这对普及标准化考试知识，深化标准化考试研究和加速传统考试的改革，诚然大有裨益。但是，我们也必须看到标准化考试热朝中的另一种倾向：对标准化考试的误解和滥用。现阶段，某些部门和学校所举行的考试，大都冠以标准化考试之名，实与传统考试无异；有些单位编制出版的试题集，也贴上了标准化试题的标签，实与试题标准化的原则、要求不符。结果造成人们对标准化考试的误解，认为凡是选择题，就是标准化试题（或只有选择题，才是标准化试题）；只要是应试者根据现成答案作答的考试，就是标准化考试（或凡能客观判分的考试，就是标准化考试）；标准化考试只能测知识，不能测能力；标准化考试的应试者无需系统掌握知识、全面发展智能，只要凭应考灵感随意打勾、涂黑点，定能稳操胜券等等。上述偏见的产生，恰恰是人们对标准化考试缺乏全面系统的了解所致。为方便社会各界熟悉标准化考试的含义、历史、现状、内容、特征、设计和实施步骤、技术手段、必备条件、功能及其与传统考试的关联和区别，我们根据现实社会的需求，针对我国目前在标准化考试的宣传和试用方面存在的一些缺陷，以及人们对传统考试和标准化考试在认识上的偏见，编著了《标准化考试的理论与实践》这本小册子，给我国中、高等学校的广大师生、考生家长、各级各类招考工作者、教育行政人员、教研人员、以及从事考试、教育测量、教育评价研究的同志，奉献一份有关标准化考试的系统资料，提供一个展望当今世界考试发展趋向和我国考试改革前景的窗口。

《标准化考试的理论与实践》一书，由五章、三个附录组成。为使读者对标准化考试的历史渊源、现阶段国内外标

标准化考试的现状有所了解，本书在第一、第五章有较为详尽的叙述。基于编著这本册子的主要目的，在于帮助读者明了标准化考试的内容，掌握设计和实施标准化考试的要领，乃至自行设计和组织实施标准化考试，故本书按标准化考试的程序，在第二、第三章中，对标准化考试的基本内容、设计步骤、实施步骤、主要特征和实施条件等，分别作了详细的介绍，各具体环节均示以范例。同时，由于我国目前正值改革传统考试、试验和倡导标准化考试时期，两种考试谁优谁劣？何者利大？何者弊小？究竟那种考试最符合中国的国情？都有待深入探讨和进一步实践。为此，我们特意编写了“标准化考试与传统考试之比较”一章，将两种考试的模式及各自的利弊，进行对比分析，以供大家研究时参考。再者，本书还编有标准化考试常用名词解释、计算公式和统计数表三个附录，便于中学、高校、招考及教研部门在革新传统考试、试验或实行标准化考试的过程中参看。

标准化考试问世于十九世纪末，它随着现代教育理论、心理理论和电子技术的发展，而不断得到改造和完善，到本世纪四十年代逐渐取代了传统考试，其使用范围日趋扩大，现已在国际上广为流行，并成为当今世界上许多国家和地区革新考试所追求的目标，也是我国高考即将采用的一种形式。近年来，我国对标准化考试展开了积极的研究和探索，在国家教委的领导下，正有计划、有步骤地进行试验，其范围和科目每年有所扩大和增加。当前，我国古老的传统考试渐趋让位，具有中国特色的标准化考试正在兴起。在这新旧考试的转折之际，下述三点，实不可忽略。

其一，标准化考试优点甚多，但并非万能，其科学性和客观性，是相对传统考试而言的，我们不能因为推行标准化

考试而全盘否定传统考试，抛弃其科学的成份。

其二，标准化考试虽已有几十年的历史，但由于科学技术的进步，社会物质生产方式的变革，它始终在变化发展之中，迄今尚无固定的模式。考试改革贵在创新，切忌把标准化考试当作僵死的公式盲目照搬，应从我国的实际出发，吸收国外确有价值的理论、技术和方法，建立中国自己的标准化考试模式。

其三，标准化考试程序严密、技术条件高、专业性强，在设计或实施标准化考试的过程中，必须严格按其程序和规则进行，要防止随意拼凑、误用或滥用。否则，将有益于考试的科学性和权威性，乃至影响我国教育体制改革根本目的的实现。

实现考试标准化，既是我国考试改革的必由之路，也是我国社会经济发展的客观要求。但是在现阶段，标准化考试毕竟还是我国考试研究和实践中的新课题，没有现成的科学理论和方法可资遵循，加之水平有限，时间仓促，资料不丰，书中错误、疏漏之处在所不免。我们是抱着以丑求美、抛砖引玉的思想，将《标准化考试的理论与实践》一书公开的。敬祈专家、读者批评指正。

书中引用的中外有关资料，是许多同道者辛勤劳动的结晶，在此特作说明，并深表谢意！

作者

一九八五年十月于武汉

目 录

绪 言

第一章 标准化考试的源流、产生与发展

第一节 传统考试的演变与标准化考试的含

义..... (1)

一 传统考试的演变..... (2)

二 标准化考试的含义..... (4)

第二节 标准化考试的源流、产生及其发

展..... (6)

一 初创阶段..... (7)

二 徘徊阶段..... (9)

三 流行阶段..... (10)

第二章 标准化考试的内容、设计与实施

第一节 标准化考试的基本内容..... (14)

一 命题标准化..... (16)

二 答案标准化..... (17)

三 施测标准化..... (18)

四 评分标准化..... (19)

五 计分标准化..... (20)

六 分数解释标准化..... (21)

第二节 标准化考试的设计步骤..... (22)

一 确立考试目的..... (23)

二 编制考试大纲..... (25)

三 编题设计..... (31)

四 编、审试题..... (34)

五	预备测验	(38)
六	制 卷	(66)
七	制定标准	(69)
八	编写考试说明书	(73)
第三节	标准化考试的实施步骤	(74)
一	制订实施方案	(75)
二	印制试卷、答卷	(76)
三	设置试场	(77)
四	施 测	(78)
五	阅卷评分	(79)
六	分数转换与统计分析	(80)
七	编写考试分析报告	(87)
八	处理考试结果	(87)
第三章	标准化考试的特征与实施条件	
第一节	标准化考试的主要特征	(89)
一	标准稳定	(89)
二	内容广博	(90)
三	质量有控	(91)
四	对象开放	(93)
五	实施独立	(93)
第二节	实施标准化考试的基本条件	(94)
一	三支队伍	(95)
二	一套机器	(98)
三	一个“仓库”	(99)
四	一套制度	(101)
第四章	标准化考试与传统考试之比较	
第一节	两种考试模式	(102)

一	传统考试模式	(102)
二	标准化考试模式	(104)
第二节	两种考试比较	(105)
一	整体比较	(105)
二	流程比较	(107)
三	利弊比较	(108)
第五章	中国标准化考试管见	
第一节	我国实行标准化考试的必要性与可能性	(111)
一	我国实行标准化考试的必要性	(111)
二	我国实行标准化考试的可能性	(114)
第二节	我国实行标准化考试的现实基础与必须创设的条件	(116)
一	我国实行标准化考试的现实基础	(116)
二	我国全面推行标准化考试必须创设的条件	(117)
附录一	标准化考试常用名词解释	(121)
附录二	标准化考试常用计算公式	(135)
附录三	标准化考试常用统计数表	(146)
表 1	正态曲线的面积 (A) 与纵线 (y)	(146)
表 2	积差相关系数 (r) 显著性临界值	(154)
表 3	等级相关系数 (r_s) 显著性临界值 (单尾检验)	(155)
表 4	弗拉南根二列相关系数 (r_{xy}) 估计数表	(156)
表 5	T 值分布	(157)
表 6	卡方 (χ^2) 值 分布	(158)

第一章

标准化考试的源流、产生与发展

第一节 传统考试的演变与

标准化考试的含义

在现代社会中，考试作为一种有意识的测度、甄别人的素质与智能个别差异的活动，已渗透于人类生产和生活的一切领域。各类社会成员，无论受过正规的学校教育与否，也不管是否亲身经历过考试的检测，他们对考试都有直接或间接的了解。就此而论，考试可谓世人无所不知。但是，对考试这种人类社会特有现象的内在规律，对其产生、发展、演变的成因，对今昔考试模式的异同与因果关联，对各种考试功能的区别，以及标准化考试的含义、内容、程序、技术手段和特征等，绝大多数人还不熟悉。在传统考试急剧变革、标准化考试广为流行的今天，什么叫标准化考试？怎样设计和实施标准化考试？它与传统考试是什么关系？两者区别何在？为什么我国要尽快实行标准化考试？这是社会各界至为关心并急需了解的问题。为便于正确理解和认识标准化考试的含义、产生与发展的状况、内容结构、设计与实施程序，以及标准化考试的社会功能，有必要回顾一下考试产生、发展的演变历史。

一、传统考试的演变

作为人类社会特有现象的考试,历史悠久,源远流长。它伴随着人类社会的演进,经历了一条极为漫长的发展道路。

早在原始社会的后期,随着畜牧业和农业的逐步分化,以及氏族和部落的形成,生产技术日趋复杂,氏族、部落之间的冲突与战争相应加剧。为了促进社会生产力水平的提高,维系氏族和部落的生存与发展,用生产、军事、图腾式宗教、氏族法律与风俗等方面的经验和技能武装青年一代,已成为社会的迫切需要。教育开始摆脱自然状态,渐渐变成一种有意识地对青年传授各种经验和技能的活动。与此相适应,一种鉴别接受经验快慢和模仿技能熟练程度高低的活动——考试,也就问世于人类。各部落和氏族开始用实际操作、演示等形式,有选择地吸收征战人员和部落正式成员。譬如,当时的男性青年人若要参加氏族或部落之间的战争,需在部落或氏族首领的监视下,当众检验其骑马、射箭、搏斗等方面的经验与技能,经过多次公开的、有时甚至是极痛苦的考验之后,才能成为被公认的军人。年满二十岁的青年,要成为氏族、部落的正式成员,取得参与各种社会活动的权力,除对风俗习惯、宗教礼仪规则等方面的实际考验外,还需举行冠礼仪式,由族长或部落首领,以及经选择参加举行冠礼的大宾,对接受冠礼的青年进行宗教礼拜方面的行为举止检验。当人类进入古代文明时代——奴隶社会和封建社会后,由于生产力和科学文化水平的不断提高,体力劳动与脑力劳动明确分工,社会随之分解为各个阶级,教育从生产劳动过程中分化出来,出现了学校这种专门的教育机构。考试因此由原来那种全民性的、开放性的、无所固定规则的

原始形态，发展成为为统治阶级服务的、封闭性的、具有一定规程的检测手段，被统治阶级用作检测、甄别、选拔所需人才的工具。在古希腊，斯巴达的农业贵族检验青少年是否符合他们要求的武士，需要经过严格而残酷的考试，如通过跳足行走、穿单衣过寒冬来测试意志；用让青年参加“埃弗比”集团的演习，在夜间包围奴隶住所，屠杀奴隶等活动，检验他们的勇敢精神。雅典的商业贵族，在文法和弦琴学校，对学习期满的学生要进行字母、缀音、书法规则、诗歌、曲谱等口头背诵测试，以及音乐舞蹈、书写等方面的演示测验；在体操学校，则通过角力、竞走、跳高、掷铁饼、投枪等各种比赛，来测量学生的学习成绩。我国商、周时期的学校教育中，普遍采用问讯、操作表演式的考试，以判定学生的学业成绩和实际本领，选拔所需人才，如《静殷》铭文就记有周厉王亲临辟雍，在灵池旁令太子静（周宣王）进行实射，测试其两月学习射箭的成绩。在奴隶社会时期，考试已成为一种有意识、有目的、有一定规则的活动，并有了测量人的素质和智能的内涵，但它尚未形成独立的概念，与教育没有分离，更谈不上形成独立体系。在此后的数百年中，考试的发展速度十分缓慢，虽考试内容在不断增加，但在形式方法上始终局限在面试范围，直到公元606年我国创立科举考试为止，考试才出现了第一个飞跃，即由面试为主过渡到笔试为主，或以两种方式测量考生的知识、能力、素质与技能，并成为—种定型的主观考试。科举考试在我国使用达1300年之久。其考试形式和方法，为我国至今所沿袭。考试的这种飞跃，在日本，因其效仿中国，于公元709年出现，而在欧洲十七世纪才得以实现，即由伟大的教育家夸美纽斯第一次提出了用考试选拔学生和评价学生成绩的观点，并得到了社会的

承认。夸美纽斯关于考试的论述和实践，将西方考试推上了一个新的阶段，为西方考试体系的形成奠定了坚实基础，在考试的发展史上立下了不可磨灭的功绩。随后，东西方不少国家相继借鉴中国考试的经验，逐步兴起了就业考试、文官选拔考试，使主观考试得以成熟和完善，形成了独立的考试体系，并广泛应用于社会生活的各个领域。根据社会的需要，通过考试甄别、评价、选拔人才，调节社会各劳动部门的人员结构和智能结构等，以发挥其独特的社会功能。到十九世纪，整个人类社会的生产力、文化教育和科学技术等有了巨大发展，社会经济结构和产业结构发生了极为深刻的变化，社会对测量人才素质与职能的考试要求更高，而传统考试由于命题主观、测试面窄、工作量大、可控性差，所测结果不能真实准确地反映应试的水平，其弊端日趋突出，与社会发展的需求愈来愈不适应。为使考试更好地服务于当时的社会，西方考试界在革新传统考试的基础上，于十九世纪创立了客观测验。它虽是主观考试的翻新和发展，但依然是一种非标准化的考试，客观测验的设计和实施，仍不能有效控制主客观因素的干扰，测得结果误差大，信度、效度不高。时至二十世纪四十年代，现代考试理论和统计学的新发展，电子计算技术的产生，为考试彻底冲破主观考试的圈子，根除客观测验的弊端，实现考试标准化奠定了科学基础，提供了有效控制测量误差的技术条件，从而出现了考试的第二次飞跃，即由非标准化客观测验进入了标准化考试的新阶段。

二、标准化考试的含义

综观考试的历史，考试随着人类社会的形成而产生，发展而演变，为一切社会所必需。因此，从宏观说，考试既是人

类社会的永恒范畴，又是人类社会的普遍范畴。它与人类社会共始终，并随社会生产和科学技术的发展而发展，不断扩大和改变其社会功能。从微观讲，考试有其自身产生、形成与发展的规律。它已经历了由不规范到规范、由面试到笔试、由手工方式到电脑控制、由主观性考试到客观性考试的演变过程，并继续处在发展变化之中。就其本质而言，考试是人类社会测度、甄别别人的素质和智能的活动。从广义上说，凡是有意识地测度、甄别别人的知识、能力和技能，考察人的思想和品德，检测人的体质的活动（不论是有组织的还是无组织的，系统的还是零碎的）都是考试。诸如：职工年终总结、鉴定，各种竞赛，文化技术考核，工资评定，科技、工程人员、教育工作者的职称评定，教育、教学质量评估，学业成绩测验，招生、招工、招干考核等等。从狭义上说，考试是根据一定社会、阶级或一定社会集团的客观要求，对应试者所进行的一种有目的、有计划、有组织的测量其知识能力、素质、技能等个别差异的活动。如教育系统和社会其他系统所进行的学业成绩考试、招生考试、就业考试、技术业务考试、升职晋级考试，等等。

迄今为止，人类所经历的考试，可分为主观性考试和客观性考试两大类。主观性考试亦即传统考试，应试者运用所学知识，根据自己的理解自由回答主试者的测试。客观性考试，系指应试者根据考试内容所提供的信息和现成答案，经过自己的分析、理解或推断进行选择，固定回答主试者的测试。

标准化考试属客观性考试，是依据现代考试理论，借助现代统计方法和电脑技术，严格按照系统的科学程序设计和实施，有效控制各种误差，具有统一标准的考试。其核心是

量尺标准、程序科学、误差有控、信效度高。

第二节 标准化考试的源流、 产生及其发展

标准化考试产生于十九世纪末二十世纪初，其历史仅八十余年，今天已风行世界。溯源追根，标准化考试是三股源流汇集而成。这三股源流是：1）德国的实验心理研究；2）英国关于心理个别差异的研究；3）法国关于诊断异常儿童智力的研究。现简介如下：

1. 实验心理学的发展对创建标准化考试的启示。实验心理学的鼻祖冯特（德国心理学家），于一八七九年在来比锡大学创设心理实验室，以实验方法着手心理研究，目的在于进一步探究心理本质问题。为便于观察，他要求实验方法必须严密，尤其要注重条件的控制，同时还要有周密的计划和测量方法的精心设计。由于采取了上述措施，冯特的实验心理学终于取得突出成绩。他的实验计划与测量方法，给考试理论和实际工作者以有益启示，并被应用于考试领域。

2. 个别差异研究对标准化考试的影响。伟大的生物学家达尔文，于一八五九年发表了科学巨著——《物种起源》，其中心论旨是进化论，阐述了关于物种各成员之间的变异问题——物种个别差异问题。此后，英国学者高尔顿（达尔文的表弟），在达尔文进化论思想的影响下，于一八八二年在伦敦设立人类学测量实验室。他通过搜集人类生理特征和心理特征的资料，并对双亲与子女、兄弟姊妹间、双生儿等的差异程度和类似程度进行研究，运用测量和统计方法，

从存在于个别差异的单纯心理过程中获得系统资料，将人类能力数量化，进而开创了心理科学新领域——心理测量。后来，它被教育借鉴，用以测试学生的能力和学习效果。从这种意义上说，它是教育测量的先声。所以人们认为，标准化考试起源于心理测量。

3. 诊断异常儿童智力的研究对标准化考试的促进。法国医学界在十九世纪展开了对智力有缺陷儿童的研究，希图探明原因，以期对有缺陷儿童有所改善。智力测验的创始人比纳，对当时巴黎一些学校中许多学习效果极差的儿童产生兴趣，并着手研究这些学生成绩低劣的原因。一八九五年他通过与安利合作，设计出一套智力测验方法，于一九〇三年发表了《智力的实验研究》。后应法国教育当局甄别智力落后学生的要求，比纳在西蒙的协助下，于一九〇五年创制了第一个适用于智力测验的量表——比纳西蒙量表。比纳认为，儿童的智能是随着儿童的成长而增加的，于是，他根据不同年龄制成了年龄常模量表，后来被称为心理年龄。由于比纳西蒙量表逐步为社会普遍应用，因而也引起了教育科学研究工作者的重视。智力测验虽与教育测量的目标各异，但原则和方法有相似之处，故对标准化考试的发展起了积极的促进作用。

以上三股源流，加上美国人事工程思想（把人类能力——量化为数目，以代表个人的价值）的促进，使标准化考试得以产生和发展。时至今日，标准化考试的发展历程大致可分为三个阶段。

一、初创阶段（一九〇四——一九三〇年）

标准化考试产生于二十世纪初，但标准化考试的观念和方法出现较早，如十九世纪六十年代美国教师菲奢编制的