

考试理论研究

于信风 著
辽宁人民出版社

KAOSHI LILUN YANJIU

考试理论研究

于信凤 著

辽宁人民出版社

1989年·沈阳

考试理论研究

Kaoshi Lilun Yanjiu

于信凤 著

辽宁人民出版社出版 辽宁省新华书店发行
(沈阳市南京街6段1里2号) 沈阳新华印刷厂印刷

字数: 190,000 开本: 850×1168 $\frac{1}{32}$ 印张: 9 插页: 2

印数: 1—1,300

1989年12月第1版

1989年12月第1次印刷

责任编辑: 李可可

版式设计: 任 和

封面设计: 赵多良

责任校对: 王绍斌

ISBN 7-205-01083-7/C·95

定价: 4.80元

前 言

人类的考试活动，是在社会政治、经济、文化发展的推动和制约下进行的。古代和现代、中国和外国（如欧美国家），考试制度及其所发挥的功能的差异，都可以从政治、经济、文化背景上找到合理的解释。当代，是改革的年代，改革的大潮席卷全球，冲击着社会的一切领域。我国经济、科技、教育、政治体制改革方案和措施的陆续出台，要求着、推动着我国考试工作进行相应的变革。稍加分析就会发现，当前，我国的考试工作面临着两大转变：一是随着国家公务员考试和各种职业（岗位）资格考试的兴起，考试将由教育测量的一种手段、教育工作的一个环节扩大为一种社会性的活动，为适应这种转变的需要，将产生出与学校教育相分离的为全社会提供考试服务的办考实体——国家兴办的或民办的考试服务机构，从而形成与学校教育系统平行的社会考试系统；二是随着考试在人才选拔、评价、录用、管理中作用的扩大，要求考试由校内的教师自编测验发展为社会性的标准化考试，从而实现传统的经验考试向现代的科学考试的转变。这两大转变，是商品经济发展的必然要求，是现代科学技术应用于考试的必然结果。

考试工作的上述转变带来了一系列实际问题和理论问题。如：考试机构的设置、建设和管理；资格考试、选拔考试和成绩考试的区别和联系；各类考试的功能和后效；考试设计和试卷编制的科学化；电子计算机在题目分析和考试管理中

的应用；等等。从理论上，可将这些问题归结为以下四个方面：考试的社会性质和功能，考试系统与社会母系统的关系问题；大规模社会性考试的管理和国家对考试机构及其活动的管理问题；不同类别考试的特点、区别和联系问题；考试设计和试卷编制的科学化问题。此外，还有考试技术问题。

关于试卷编制和分数解释，教育测量学已提供了很好的理论基础，在这方面，由于我们的研究中断了三、四十年，现在的主要工作是“引进”，是“消化”，是在考试实践中的应用和发展。但是，测试编制理论并没有回答考试的性质和功能问题，没有解决考试管理问题，对它的研究也不能代替对不同类别考试的专门研究。就是说，考试理论并不就是测试编制理论，考试学并不就是教育测量学。就象教育学首先要论述教育与社会的发展、人的发展的关系，在阐明教育的一般规律后还要专门研究智育、德育、体育、美育的特殊规律，也要探讨教育管理，为教育分支学科提供理论基础一样，在考试从教学过程中分离出来成为一种社会性活动后，作为指导考试实践的考试学也应研究考试与社会发展的关系，研究不同类别考试的特殊规律，研究考试管理，为考试不同方面的专题研究提供理论基础。建立这样的考试学，是考试工作发展的需要，是来自于实践的理论课题。

基于此，自1982年以来，作者多次提出加强考试科学研究、建立考试学的问题，并把1985年末、1986年初，在我即将离开考试组织管理工作之时撰写的实际工作经验总结和理论探讨心得，斗胆定名为《考试学引论》（辽宁人民出版社1987年1月出版。下简称《引论》）。《引论》重在“引”

上，希望引起考试理论工作者和实际工作者的关注，引出“考试学”的系统理论。写作《引论》的当时，我正被考试组织中的许多实际问题困扰着，因此，书中的结构安排和内容选择，强调了对实际工作的针对性和直接适用性，对于考试工作组织者关系不直接的理论探讨，希望留待我转到科研部门之后再抽空进行。1988年上半年，我便按考试理论至少包括四个方面的上述构想，写了四篇十二章，把我对考试学的对象和理论体系的意见作为绪论，拟定名为《考试学原理》，于6月中旬完稿。

由于理论界和出版界众所周知的原因，现在是“写书难，出书更难。”1989年3月，就在我出书无门的时候，辽宁人民出版社的领导和新学科编辑室的同志们，不计利润的得失，表示愿出此书，使我本已冷却的心又热了起来。我考虑再三，实在不敢说本书构筑起了考试学的骨架，建立了考试学的理论体系，所以决定更名为《考试理论研究》；但为了与考试界的朋友交流对考试学体系的看法，就把绪论中关于考试学原理的议论保留了下来。

已出版的《考试学引论》和这本《考试理论研究》（下简称《研究》），是对象相同，但角度、内容不同的“姊妹篇”。前者侧重于“怎么做”，带有较多的描述性，后者侧重于“那么做”的理论依据，带有较多的论证性；前者强调“可操作性”，后者突出原理和方法；即使谈论同一个题目，两者的广度和深度也都大不相同。譬如，《引论》中就已提出了考试的基本功能，但它侧重于诸种功能的描述及在考试实践中怎样发挥这些功能，而《研究》则在分析考试的

社会性质和测量特点之后，着重论述考试具有诸种功能的原因；《引论》中也分析了考试的信度和效度，但那里是把它仅仅看作检验考试质量的指标提出的，并侧重于它们的实际估计，而《研究》则是在测试误差分析之后，把它们作为一种测试理论阐述的，同时还介绍了新的测试理论——题目反应理论；《引论》也讲述了考试设计，但它仅限于说明考试设计的任务、内容和工作过程，而《研究》则着重介绍考试设计的原理和方法；《引论》也分析了成绩考试、招生考试和自学考试，但它是对我国当前最重要的三种具体考试制度的描述和改革的探讨，而《研究》则从功能和测量理论方面论述成绩考试、选拔考试、资格考试三大考试类别的特点和关系；《引论》也讲述了考试的组织管理，但它仅限于管理过程的描述，而研究则着重论述作为管理基础的考试心理及其调控和考试管理原理，而不描述具体管理过程。总之，《研究》是《引论》的理论深入和发展，而不是《引论》的扩充或缩编。

在辽宁人民出版社的支持下，《考试理论研究》就要出版了。此时，我反倒有些惶恐和不安：它能给考试理论工作者和实际工作者些许的启迪吗？它能给未来的考试学大厦增添几块有用的砖石和瓦片吗？它真值得辽宁人民出版社朋友们的厚爱吗？

真的，我没有把握。

我期待着读者的检验，批评指正。

作者

1983年3月

目 录

前 言

绪论 考试学的对象、内容和方法	1
第一节 考试学的对象	1
一 建立考试学任务的提出	1
二 考试学的研究对象	2
三 考试学和教育测量学	3
第二节 考试学的内容体系	6
一 考试学的基本体系	6
二 考试学原理的主要内容	7
第三节 考试学研究方法	8
一 考试学研究方法概说	8
二 考试学常用的研究方法	10

第一篇 考试的性质和功能

第一章 考试的社会性质	14
第一节 考试的产生和发展	14
一 考试的产生	14
二 古代考试的发展	16
三 心理测验与教育测验的兴起	20
四 现代考试的特点	22

第二节	考试与社会的关系	25
一	政治、经济、文化对考试的决定作用	25
二	考试对政治、经济、文化的反作用	29
第二章	考试的测量性质	33
第一节	知识、技能和能力测量概述	33
一	测量的定义	33
二	知识、技能、能力及其可测性	35
三	能力测量的特殊性	36
四	知识、技能和能力的测量方法	39
第二节	考试的测量属性	40
一	考试的概念	40
二	考试法测量的是什么	44
三	考试法与其他测量法的比较	47
第三节	考试的测量特点	48
一	直接测量和间接测量	48
二	全体测量和抽样测量	50
三	绝对测量和相对测量	51
四	等级测量和等距测量	52
第三章	考试的功能和作用	55
第一节	考试的功能	55
一	研究考试功能的意义	55
二	研究考试功能的依据	56
三	考试的基本功能	58
第二节	考试的作用	65
一	考试的作用概述	65
二	考试的局限性	68
三	考试在教育评价中的应用	70

四 考试在人才选拔中的应用	72
---------------------	----

第二篇 考试设计和测试编制理论

第四章 考试设计的原理和方法	75
第一节 考试设计的任务和意义	75
一 考试设计的任务	75
二 考试设计的意义	76
第二节 考试目标的规定	79
一 考试目标的含义	79
二 能力结构理论简介	80
三 教育目标 (认知领域) 分类简介	83
第三节 考试内容的确定	88
一 考试内容的含义	88
二 回归分析在确定考试内容中的应用	90
第四节 考试标准的制定	93
一 考试标准的含义	93
二 多重删割法	94
三 组合分数的百分比通过法	95
第五节 试题形式的选择	98
一 主观性试题和客观性试题	98
二 主观性试题和客观性试题的类型	101
三 试题类型的选用	104
第六节 命题蓝图的编制	107
一 命题蓝图的内容构成	107
二 试题数量及其分布的确定方法	108
第五章 考试误差的分析基础	114

第一节 考试误差的概念	114
一 研究考试误差的意义	114
二 考试误差的定义	115
三 相对误差和引用误差	116
第二节 考试误差的来源	117
一 方法误差	117
二 工具误差	118
三 测试过程误差	119
四 考生临场发挥误差	120
第三节 考试误差的表现形式	121
一 系统误差、随机误差和粗大误差	121
二 精密度、正确度、准确度	125
第四节 随机误差的统计性质	126
一 随机误差的概率分布	126
二 随机误差的特点	129
第六章 考试的信度理论	131
第一节 信度的概念	131
一 信度的意义	131
二 信度的理论定义	133
三 测试的标准误差	135
第二节 信度的估计方法	137
一 稳定系数、等值系数和内在一致性系数	137
二 内在一致性系数的估计	139
三 估计信度的三类方法的应用	141
四 测量标准误差的估计	143
第三节 影响信度的因素	144
一 考生之间的差异	144

二	测试的长度	145
三	测试的难度与区分度	146
第四节	试题的分析	147
一	试题的难度分析	147
二	试题的区分度分析	149
三	试题间的相关分析	150
第七章	考试的效度理论	152
第一节	效度的概念	152
一	效度的理论定义	152
二	效度的意义和类型	153
三	效度和信度的关系	157
第二节	效度的检验方法	158
一	效度资料的搜集	158
二	相关分析方法	159
三	因素分析原理	162
第三节	影响效度的因素	167
一	考试误差	167
二	测试的长度	167
三	考试的异质性	168
四	考试的组合	168
第八章	题目反应理论	170
第一节	经典理论的局限	170
一	题目和试卷参数对于考生样本的依赖性	171
二	难于适应目标参照测验的编制	172
三	考生对答案猜测及其矫正理论的缺陷	174
第二节	题目反应函数和信息函数	175
一	潜在特质空间	175

二 题目反应函数和题目特征曲线	176
三 题目信息和测验信息	178
第三节 题目参数及其估计	180
一 题目参数	180
二 题目参数的估计	181
第四节 题目反应理论的优越性	182
一 题目参数不变性	182
二 区分度参数和难度参数的互相独立性	183
三 难度参数与能力参数度量单位的一致性	184

第三篇 考试的分类研究

第九章 考试的方法和类别	185
第一节 考试的方法	185
一 口试法	185
二 表演操作法	187
三 笔试法	188
四 典型问题处理法	189
第二节 考试的类别	190
一 考试类别的划分	190
二 目标参照考试和常模参照考试	193
第三节 标准化考试	196
一 标准化考试的概念	196
二 标准化考试的要求	199
三 标准化考试的评价	200
第十章 成绩考试、资格考试和选拔考试	203
第一节 成绩考试	203

一	成绩考试的制度和功能	203
二	成绩考试的特点和实施要求	205
三	成绩考试的分数解释	207
第二节 资格考试		212
一	资格考试制度和功能	212
二	资格考试的特点和要求	214
三	题目反应理论在资格考试中的应用	217
第三节 选拔考试		218
一	选拔考试制度和功能	218
二	选拔考试的特点和要求	220
三	选拔考试和资格考试关系的再讨论	222
四	预测效度和临界分数对选拔正确率的影响	224
五	选拔考试分数的解释	230

第四篇 考试管理

第十一章 考试心理		234
第一节 考试心理概述		234
一	考试心理的含义	234
二	考试心理对考试过程和结果的影响	236
第二节 主考心理及其调控		239
一	主考的心理行为	239
二	主考心理的调控	241
第三节 考生应考心理及其调控		244
一	动机与期望	244
二	情绪与焦虑	247
三	注意	250

四 意志	251
第四节 反馈心理及其调控	252
一 考试反馈心理	252
二 反馈心理的调控	253
第十二章 考试管理	255
第一节 考试管理体制和组织机构	255
一 宏观管理与微观管理	255
二 考试管理体制	256
三 考试组织机构	259
第二节 考试管理过程	261
一 考试管理过程的特点	261
二 考试管理的决策过程	262
三 考试管理的执行过程	265
四 考试管理的评价过程	266
第三节 考试管理原则和方法	268
一 考试管理原则	268
二 考试管理方法	269
主要参考文献	272

绪论 考试学的对象、内容和方法

第一节 考试学的对象

一 建立考试学任务的提出

人类的考试活动已有几千年的历史了，但直到本世纪初才出现第一个关于考试的理论——教育测量学的经典理论。不过，它仅仅把考试作为一种测试手段，研究其在教育效果测量中的应用，实际上它主要是一种试卷编制的理论（信度和效度理论）和分数解释的理论（量表理论）。早在30年代美国批判教育测验运动的“八年研究”，结果就已表明，考试的效果并不完全取决于试卷的编制。60年代至80年代，随着潜在特质理论特别是题目反应理论的逐步成熟和电子计算机在试题分析中的广泛应用，测验编制理论和编制技术发展到了一个新的阶段，但批评考试的声浪丝毫没有减弱。

我国“十年动乱”中，取消了考试，造成了教育的大倒退和各行业人员素质的大下降。粉碎“四人帮”之后，恢复了高考制度和学校成绩考试制度，又建立了自学考试制度，会计师、律师、教师等资格考试正在兴起，国家公务员选拔考试制度也在酝酿中，考试在人才培养、发现、评价、选拔

和管理中的作用越来越大。但同时，“分数贬值”、“高低能”、“加剧了文凭热”、“逼使青少年陷入无边的题海中”等批评声也随之而起。

在当代，考试应用得如此广泛，考试的触角几乎已延伸到人类存在的任何处所；几乎每个成年人都在考场上“身经百战”过，每个成长中的青少年都在考试的“阴影”下辛苦奋斗着。考试是实现人生价值的阶梯，考试是驱动人们的指挥棒。考试使许多人获得了成功的喜悦，同时又夺去了更多人的欢乐。人类发明了考试，又并非情愿地被考试驱使着。这一切，不能不引起人们的深思：如何评价考试的“功过”？它的“功过”全部来源于试卷的编制技术和质量吗？作为指导人类考试活动的考试理论，除了考试的测量特性和试卷的编制外，难道不应该研究考试活动的社会性质和功能、制度和政策、机构和管理吗？于是，到了80年代初，人们便顺理成章地把建立考试学的任务提到了考试研究的日程，并进行了许多有益的探索。

二 考试学的研究对象

考试学是一门研究考试现象及其规律的社会科学。它的研究对象是测度人的知识和能力个体差异的社会活动，即考试活动；它的任务是揭示考试活动的客观规律，阐述适应社会政治、经济、文化发展需要、符合考试规律的考试理论，以指导考试实践。

随着科学技术的发展及其广泛应用，社会各行各业对人的知识和能力素质要求越来越高，考试的运用也越加广泛。