

天承運

奉

刘芑

LUPENG KAOSHI WENJI

考试文集

人民教育出版社

二百七十三名第一甲
賜進士及第第二甲
賜進士出身第三甲
賜同進士出身故茲
誥示



刘芑

LIUPENG KAOSHI WENJI

考试文集

人民教育出版社



图书在版编目 (CIP) 数据

刘芑考试文集/刘芑著. —北京: 人民教育出版社, 2012. 5
ISBN 978-7-107-24114-7

I. ①刘… II. ①刘… III. ①高等学校—入学考试—文集
IV. ①G647.32-53

中国版本图书馆 CIP 数据核字 (2012) 第 038113 号

人民教育出版社出版发行

网址: <http://www.pep.com.cn>

人民教育出版社印刷厂印装 全国新华书店经销

2012 年 5 月第 1 版 2012 年 6 月第 1 次印刷

开本: 787 毫米×1 092 毫米 1/16 印张: 39

字数: 492 千字 印数: 0 001~1 000 册

定价: 52.60 元

如发现印、装质量问题, 影响阅读, 请与本社出版科联系调换。

(联系地址: 北京市海淀区中关村南大街 17 号院 1 号楼 邮编: 100081)

自序

本人在教育考试机构供职多年，始终秉持“求贤济理，询事考言，务取由衷，以观深识”之古训品评人物，自以为深谙取人之道，而事实却一再证明，其实我才是那个永远的应试者，因为没有一次对他人的考试不是对自己更严苛的考试。

通常人们会认为，提出一个问题比回答一个问题要容易，其实不然。举个例子：母亲给不在身边的儿子打电话：“儿子，你想我吗？”儿子：“当然想啊！”母亲：“你想我哪儿啊？”儿子：“想你整（个）人呗。”孩子的第二个回答显然未能对应母亲所问之意，诧异之余，母亲豁然了悟是自己的问话出了毛病。在这个事例中，是母亲的提问造成了不是她所期待的回答。我猜母亲想问的意思是：作为你的母亲，我哪方面最让你想念？如果母亲在提问时考虑到孩子的理解取向，就能避免所答非所问。这个事例同时告诉我们：沟通的困难，往往不是不知道怎样回答，而是不知道怎样提问。有时候，一个“错误”的回答恰恰指出了一个错误的提问。日常生活尚且如此，那么考试更是把“问”与“答”的对应关系推向极致。

考试其实是一个包括命题和考试实施等环节在内的评价体制，而命题则突出地反映了考试的意志、品

质，集中地表现了考试的社会文化价值和测量学意义，就此义而言，考试又是一个复杂的、动态的学术管理体系，它是由考试的知识范畴、能力目标、测试标准和测试方式以及受试信息等子系统组成。考试要根据测试目的和受试者的心智水平、学业规格等因素，构建以知识内容为载体、以能力目标为核心、以测试标准为量尺的问题蓝图。这个问题蓝图首先必须知道，其问题设计能否引导受试者进入问题情境，即问题与问题对象的性质相对应；而它又必须最终知道，通过问题蓝图，是否对受试者进行了合理区分，即问题与问题对象的结构相对应。如果说问题蓝图是编织的渔网，入网的就应该是鱼，若是泥鳅，便说明网有问题。换做学术语言表述就是：考试的终极目标是实现考试的有效性。构成考试体系中的任何一个子系统的失真或仅仅是部分的失真，都会导致考试的失效，而考试的失效又将致使我们对诸多问题的误读和误判，甚至可能引发错误政策出台，“后果很严重”。因此严肃一点说，考试也有理想，考试的理想就是追求高效度。高效度是考试的核心价值，是指导考试实践的重要理念，也是考试内容与形式改革的永恒主题。考试的学术管理归结到一点，就是对考试效度的管理，而效度概念就是放大的“问”与“答”的对应关系。

至此，我们或许可以得出这样的结论：提出一个具有考试意义或测量价值的问题，并不比回答这个问题更容易，甚至更困难。事实上，考试的本质是对人的心理特质的描述，是一种心理活动对另一种心理活动的审视与评价，这种心理互动关系本身就会带来某种天然的“误差”，物理测量的所谓“测不准理论”同样存在于对人的测量当中，考试只是最大限度地追求那个近似值，实现考试效度的最大化。因此，“学会提问”是对提问者永远的考试，甚至是更严苛的考试。

考试可能会得满分，但是命题者永远不可能找到证据证明自己的问题蓝图没有瑕疵，永远都在不断地修订问题蓝图的过程中，追逐那个越来越近的测量目标，这既是考试的遗憾，也是考试的魅力，因此有人把考试称作遗憾的艺术。从事这个工作久了就会发现，提问确是一种没有止境的艺术追求。本人过去在学校教书，向学生提问、给学生考试是经常的事，因此我不怀疑自己是否会提问题，但说实话，现在怀疑了，因为每每有数据显示某一问题设计的不足或偏颇，提示你必须检视自己的设问以及与之相关的试题设计因素。因此，根据考试对象、考试内容和试题制作等变量，对问题蓝图进行动态管理，是获得测量学参数、维护题库运作、保证考试实施的基础性工作。另外在大规模选拔性考试的激烈角逐中，每一个设问都承担着千钧分量，每一个分值都“价值连城”。一问不当，“死伤”一片；一分之差，高下失序，这种压力也迫使自己经常反躬自问。总之，一旦涉足提问，回报便是来自外部和内部的不计其数的诘问和反问，这就是环绕考试而产生的一种特殊话语氛围，这个话语氛围实际上是一个取之不尽、用之不竭的思维宝藏和学术宝藏。作为“当事人”，只有在不断地研究、领会和回应这些诘问和反问的过程中，才能获得对提问的深刻认识，提高试题设计能力，并逐渐形成以准确、简约、透彻为主要特征的提问思维品质。可以说，提问即是对提问者的提问和追问，同时也反映提问者的能力和智慧、修养与境界。

从这个意义上说，本书的内容似可归结为一个主题：“问题”是怎样产生的。

当然，考试不仅是一门测量科学，还是一种文化现象，尤其是在我们这样一个具有悠久考试历史的国家，人们大都不能逸出考试文化的笼罩。事实上，考

试是国人的一个心结，萦绕千载，幻化了无数的历史故事；考试又是国人的一个情结，千年演绎，创造了炫丽的文化景观，遍数古制，没有哪一种世俗社会的文化建制能够把个人命运和国家命运联系得如此紧密。然而自从有了考试之后，对考试的争议就从来没有停止过，我们和考试之间的“恩怨”，也并未因社会的变迁稍有消解，反而历久弥深。如果说从古到今，我们对什么又爱又恨、荣辱共之，屡废屡兴、欲罢不能，大概就只有考试了，实可谓千年一叹。诚然，古今之社会性质和考试性质截然不同，但传统观念仍可在不同的环境里利用不同的材料继续编故事。当代考试行为与传统考试文化的碰撞，带来了一些新的社会现象和心理现象，也为我们继承考试的历史文化遗产提出了新问题，对此，书中亦稍有涉及，散见于不同的篇目之中。

本书分为三部分。第一部分是“考试研究”，主要是对考试一般性问题的探讨。从20世纪90年代迄今，我国教育考试进入了一个深化改革的时期，因此这一部分涉及考试和命题改革的内容比较多。第二部分是“历史学科考试研究”，主要是对历史学科命题理论与实践的研究，这与本人从事的具体业务工作有关，是本书的主体内容。第三部分是“考试时议”，这部分收录的是2003至2009年期间为《中国考试》杂志撰写的卷首语，多属本人对考试现象的有感而作，也有一些对考生的励志性文字，此番经删削选录，可以从中一窥当时的考试情势，也算是对当代考试文化的某种著录吧。

和很多人一样，本人最烦的事情就是考试，然而世事弄人，在年届不惑之时“误入”考试之门，成了让别人最烦的人。孰不知，考别人比自己考更闹心，因为我必须无休止地把“是什么？”转换成“为什

么?”、把“你想我哪儿啊?”转换成“你应该怎么想啊?”再抛给别人。这实在是一个绞脑汁、榨精力、耗心思、熬体力，成天感觉头脑不够用的活儿，而且再有成倍的力量投入进去，也会被这个“黑洞”吞噬，即便如此，错误和疏漏仍如影随形。一个突出的感受是，在营造命题语境和试题情境的时候，自己虽然心里明白，可转译成语言总是与内心领悟有很大距离，并且这个距离似乎难以消弭，这也是我无法满意任何一次考试的根本原因。我有时甚至想，语言这玩意本来就是靠不住的，而缺乏沟通能力的语言又是更靠不住的，如果想用一种语言表达对应可能出现的各种语言表达是最靠不住的。每逢此时，一种莫名的无奈感便油然而生。本人在此略抒衷肠，绝无博取同情的企图，也没有对考试失去希望，更不是宣称任何语言都不能相信，我只是想说，本书仅是我从事考试工作以来的一个学习记录，或曰应时之作，内容驳杂而浅尝辄止，学识不足而走笔匆匆，和这个“遗憾的艺术”一样，书中所论远未反映命题的理论与实践以及考试文化内涵的应有之意，其错误、疏漏和一些现在看来近乎陈腐的内容在所难免，对此本人一律没作改动，一是保留历史痕迹，二是冀望识者指教。

刘 芑

二〇一二年元月于北京溪翁山庄

目 录

自序	/ 1
----	-----

第一部分 考试研究 / 1

论考试 / 3
论考试内容 / 7
考试内容考辨 / 14
论试题的合理性 / 18
考试内容与考试方法 / 24
现代教育测量启示 / 29
论综合考试与综合能力 / 38
考试社会学刍议 / 45
考试行为与市场机制 / 51
标准参照考试在自考命题中的应用 / 56
从命题谈培养自学能力 / 61
全国自考文科命题思路 / 65
高考命题与素质教育 / 69

第二部分 历史学科考试研究 / 75

历史事实与历史思维 / 77
历史思维的发生
——兼论历史思维的测量 / 83
历史思维评价 / 92
测量视域中的历史 / 98
历史教育与命题 / 105

历史学科的题型功能 / 116
态度、历史态度及命题理念纲要 / 125
历史学科能力因素分析及题型评析 / 135
历史学科考试测量的理论与实践 / 146
历史学科考核内容与要求 / 185
历史学科教育测量概论 / 218
历史学科的评价目标 / 237
历史高考考试内容与历史教学 / 330
历史试题的编制 / 337
历史学科测量的评价 / 379

第三部分 考试时议 / 407

新年第一题 / 409
考试的传说 传说的启示 / 411
黄金法则 / 413
星空 / 415
走进高三 / 417
秋思 / 419
走出悖论 / 420
考试立法
——对考试的考试 / 422
天下智仓 / 424
理性之光 / 426
生命的祭献 / 428
高原诗笺 / 430
散议舞弊 / 432
“题目”作文 / 434
蚁行 / 436
考试改革与考试安全 / 438
路边的银杏树 / 440
人生如“画” / 442
百年期盼 / 444
春雨如诗 / 446

激情不再? / 448
对自己要公平 / 450
精神苦旅 / 452
考试时疫 / 454
知识的遭遇 / 456
考试科学散论 / 458
应试之累 / 460
鹰赋 / 462
考试内容简议 / 463
下雨了/这真奇妙 / 465
试题的本质 / 467
红 / 469
考试的起源和考试的缘起 / 471
文凭——“入世”的请柬 / 473
有一种成长 / 475
教你应试 / 476
了犹未了 / 478
说话与作文 / 480
三十年河西? / 482
考试的“哲学思考” / 484
自学考试与学业评价 / 486
解放考与解放人 / 488
走进大学 / 490
走近大师 / 492
学习型社会和考试型社会 / 494
世纪之爱 / 496
创造著名考试 / 498
事实·价值·评价 / 500
我们需要什么样的考试 / 502
古训 / 504
希望在冬季 / 506
满天言语 / 508
考试的隐性功能 / 510
游山有感 / 512

高考之高 / 513
什么是知识 / 516
牧笛中的考试 / 518
大考之最 / 520
我的假日酒店 / 521
考试技术 / 522
测不准的诱惑 / 524
旧题重温 / 526
夏天的记忆 / 529
试题评价与考试评价 / 530
我们“复兴”什么 / 532
远山的回应 / 534
安全之道 / 535
三十而“立” / 537
心中的甲骨文 / 539
试题的观赏性 / 541
优美感和崇高感 / 543
考试内容异说 / 545
告别 / 547
“提裤” / 549
“重要性”之议 / 551
社会转型与考试转型 / 553
“悲剧”的冷艳 / 555
考试兴衰和大国沉浮 / 557
人性化考试 / 560
力 / 562
大漠孤泓 / 564
榨题 / 566
心之官 / 568
试题“保鲜” / 570
唤醒标准分 / 572
经济全球化和考试国际化 / 574
考试嬗变和文化认同 / 576
罢黜百家，独尊考术 / 579

考试文化的反思与重构 / 581

难度的故事 / 584

教育·考试 / 587

信度信否? / 589

“现代性”问题 / 591

功名与公平 / 594

金榜余晖 / 597

豪华危机 / 599

水到而渠未成 / 601

学会思考 / 603

青山着意 / 605

大国的困惑 / 607

第一部分

考 试 研 究

论考试

在现实生活的诸多困惑里，也许考试会居于其中。这是因为，如果没有考试，人们就会制造很多考试来区分资质、品评等第以建立人才秩序。而一旦有了考试，尤其是考试成为主要的控制力量的时候，人们又会感到很不舒服，意欲摘掉这个“紧箍咒”。我国历史上考试屡兴屡废、沉浮不已的事实就反映了这种困惑。

从考试现象的发生、发展、演变的历史来看，只要存在社会分工，考试就是不可避免的。因为从本质上说，考试是通过各种评价手段规定或转换人的社会角色，只要人们需要对自己充任的角色取得社会性确认，就要通过这种评价。历史上所谓的取消考试，事实上也只是变换了一种评价方式而已。

既然考试是客观需要，何以又屡遭废止呢？关键是考试的操作与社会的运行发生了难以协调的矛盾。我们知道，社会事物都具有历史范畴的属性，考试也不例外。这就是说，考试产生于某种特定的社会历史环境，为当时的社会需求服务。当历史发展了，社会变化了，考试的生命亦将面临终结，此时若考试固守其旧，人们便会强迫它终结，若考试顺时而变，服务于新政，体现时代的要求和价值观念，人们将会承认其存在。

那么，考试如何顺乎时势、图新以进取呢？

其一，建立新的公平观

公平性是考试的立身之本，但考试的公平性是相对的，它只能体现一定历史条件下的公平法则，因而其公平性又是暂时的。比如中国历史上魏晋时期的九品中正选官制度，对于士大夫来说是公平的，但对于庶族阶层来说就是不公平的。九品中正寿终正寝之后，随即产生了科举制度。科举制把中国封建社会考试的公平性发挥到了极致，但它对于无法获得教育资源的最基层的民众来说还是不公

平的。由此可见，考试的公平性是受历史条件制约的。根据变化了的社会环境和历史条件，建立新的公平观，并在此基础上对考试进行改造，是考试谋求自身发展的必然道路。

我国的改革开放和“科教兴国”战略决策的实施，为以提高国民素质为宗旨的教育事业创造了良好的社会环境和政策背景。考试作为教育事业的重要内容，应该十分清醒地意识到自身的责任，要开发尽可能多的考试品种和创造周到的考试的社会服务机制，为人们创造多途径受教育的机会和提供社会需求以及实现个人价值的机会，从而形成最广泛的社会公平。这是我国考试界面临的改革和发展的机遇。

其二，科学认识和合理运用考试的社会控制功能

从社会学的角度审视考试，考试显然具有明显的社会控制功能。概括来说，考试的社会控制功能就是考试制度和考试程序作用于人的观念和行为的过程。有什么样的考试制度，就会有什么样的考试观念。比如我们现行的教育升学考试制度，它使人们建立了科学文化知识水平是体现自身价值和社会价值重要标志的观念。而“文化大革命”期间实行的推荐上大学的制度，使人们趋向于科学文化知识不是体现自身价值和社会价值的观念。同样，有什么样的考试程序，就会有什么样的考试行为。比如高等教育自学考试，其考试程序的主要特征是长时间的、滚动式的课程自学考试，这种以考促学的形式既方便了考生，又极大地保护了、鼓励了自学者学习热情，形成了学习与生产实践相结合的学习行为和踊跃投身自学考试的考试行为。可见，考试制度和考试程序对人们的观念和行为具有相当强的引导作用，并通过这种引导实现对社会的有效控制。由于考试无处不在，人们几乎不可能完全抛开考试这个因素去设计自身的发展、个人社会角色的转变和参与社会活动。因此，正确理解和运用考试的社会控制功能，是深入发掘人才潜力、科学配置人才资源、合理疏导人才流向、创造良好的社会人才生态环境的重要研究课题。

应该说，我们当前的考试尤其是教育考试制度和考试程序还存在有待进一步完善的问题。比如大学入学考试，高等学校招生自主