

“教育治理与学校变革”研究丛书

华东师范大学教育管理学系组编

考试评价论

张远增◎著

K A O S H I
P I N G J I A
L U N



华东师范大学出版社

“教育治理与学校变革”研究丛书

华东师范大学教育管理学系组编

考试评价论

张远增◎著

K A O S H I
P I N G J I A
L U N



华东师范大学出版社

图书在版编目(CIP)数据

考试评价论/张远增著. —上海:华东师范大学出版社,
2018

(教育治理与学校变革研究)

ISBN 978-7-5675-8380-1

I. ①考… II. ①张… III. ①考试制度—教育评估—研
究 IV. ①G424.74

中国版本图书馆 CIP 数据核字(2018)第 222482 号

考试评价论

著 者 张远增
策划组稿 李文革
项目编辑 师 文
审读编辑 李文革
责任校对 时东明
装帧设计 庄玉侠

出版发行 华东师范大学出版社
社 址 上海市中山北路 3663 号 邮编 200062
网 址 www.ecnupress.com.cn
电 话 021-60821666 行政传真 021-62572105
客服电话 021-62865537 门市(邮购)电话 021-62869887
地 址 上海市中山北路 3663 号华东师范大学校内先锋路口
网 店 <http://hdsdcbs.tmall.com>

印 刷 者 常熟高专印刷有限公司
开 本 787×1092 16 开
印 张 17.75
字 数 277 千字
版 次 2018 年 10 月第 1 版
印 次 2018 年 10 月第 1 次
书 号 ISBN 978-7-5675-8380-1/G·11536
定 价 48.00 元

出 版 人 王 熠

(如发现本版图书有印订质量问题,请寄回本社客服中心调换或电话 021-62865537 联系)

丛书编委会

主 编 李伟胜

编委会 李伟胜 魏志春 张万朋
胡耀宗 朱军文 胡东芳
董 辉

《“教育治理与学校变革”研究丛书》序言

有个秘密，等着我们一起去探究。这个想法让我们充满好奇，让我们已有的学术探索之旅多了一种意义，让我们开始了新的尝试，包括组织出版这套丛书。

这个秘密，就是上海教育“秘密”或中国教育“秘密”。当然，延伸开来，这样的“秘密”不只涉及一个地方、一个国家的教育系统，因而具有更广泛的国际意义。自从上海学生于2009年开始在三年一次的PISA(Program for International Student Assessment, 由OECD组织实施)测试中取得佳绩之后，这个“秘密”就已经吸引了诸多同行。后来，一系列的调查、研究，尝试从学生学习范式、教师教学方式、教师专业发展方式等方面解释这个“秘密”；上海教师参与的2015年TALIS(The OECD Teaching and Learning International Survey, 由OECD组织实施)，让这个“秘密”得到了更多解释……不过，已有的研究，似乎更多强调从学生个体、教师个体的角度来展开探究，较少从师生所在的群体（如上海初中稳定的班级生活、上海教师常态化的教研团队）的角度开展探究；至于从更上位的学校（作为教育组织）和教育系统的角度来作出探索，更为少见。显然，这是一个很有意思的探索空间。

对于身处上海这片教改热土、持续多年研究教育管理的学术团队来说，这似乎是一个需要直面的摆在“家门口”的“秘密”——从教育治理和学校变革的角度来看，如何解释上海教育“秘密”？我们很好奇，国际国内同行也很好奇——大家常在一起追问这个问题，尝试解谜，并且兴致渐浓，因为这不仅涉及如何看待上海教育、中国教育，还涉及如何在全球视野中理解不同文化生态中的教育系统，理解其中的教育组织和发展主体。

恰在此时,华东师范大学教育管理学系所在的学术团队正在凝神聚力关注“教育治理与学校变革”研究,以便在参与“双一流”建设(建设世界一流大学和一流学科)和教育学“高峰学科”建设的战略行动中拓展新的发展空间。自2000年建系以来,通过持续的探索,我们这个团队重点从国家、区域和学校这三个层面关注教育系统运行,在与之相关的教育领导与行政、学校管理与改进、教育评价与测量、教育人力资源开发、教育法规与政策、教育经济与财政、教育资源配置及使用等方面有了一定的学术积累。现在,我们可以尝试聚焦教育系统变革,特别是在“教育治理”这个领域,研究区域性的教育治理体系如何促进学校变革、进而促成学校成员(校长、教师和学生)的发展。

经过一段时间的初步探索,我们合作制定了团队发展规划并由此为团队成员的发展搭建平台、提供支持,举办了“教育治理与学校变革”国际研讨会并由此展开与国际同行的交流合作。在此基础上,我们商定:组织出版一套“教育治理与学校变革”研究丛书,从而将探究上海教育“秘密”的具体行动纳入更开阔的国际学术平台上,以便促进这一领域的系统研究,力争在国际学术视野中作出一点新的探索。

这套丛书,由华东师范大学教育管理学系组织策划,主要用来出版“教育治理与学校变革”领域的成果。鉴于相关的研究视野很开阔,研究者可以选择多元的视角或主题,本丛书将秉持开放的态度,欢迎更多研究者将自己的优秀学术成果纳入到丛书出版计划之中。

同时,鉴于相关研究尚有许多新空间需要拓展,我们真诚地期待更多同行关注、支持我们的努力,并提供宝贵的指导意见或建议,以便让我们一起参与的学术研究和教育事业有美好的未来。

让我们一起充满希望地探索,着眼千里,渐积跬步。

李伟胜

2018年6月6日

考试评价：考试研究的新领域^①

——代前言

一、何谓考试评价

在考试实践中经常会面临如下的问题：

问题 1：在个人或团体组织实施某种自己需要的考试时，他们需要判明所实施的考试是否满足自己的需要，考试自身是否合乎测量规律，考试本身有何改进余地及改进方法等；

问题 2：在个人或团体作为考试对象参加他人按自己意愿设计的考试时，他们需要判明自己所参加的考试是否体现了自己的要求，考试所提供的问题情境是否有利于自己通过解答问题了解自己的真实状态，考试自身的设计是否合乎测量规律，考试本身有何改进余地及改进方法等；

问题 3：在个人或团体委托他人按照自己的要求对特定的对象设计、组织、实施考试时，他们需要判明被委托者所举行的考试是否合乎自己的需要，考试自身的设计是否合乎测量规律，考试本身有何改进余地及改进方法等。

以上三个问题均涉及对考试的价值判断问题，只利用缺乏价值判断的考试质量分析方法是难以解决它们的，要有效解决这三个问题，必须运用评价方法来研究考试。

所谓考试评价就是指对特定考试之合乎其对应的特定价值主体（指对该考试的占有者、决定者和发起者，本文将其称为选用考试者）需要和合乎有关考试

^① 原文载《考试研究》2005 年第 1 期，第 5—11 页。

规律(测量规律)的程度进行判断的活动。考试评价以考试作为评价对象,目的是判断考试之合乎目的性和合乎规律性,为选用考试者正确使用考试结果提供决策依据,为进一步改进和完善考试提供决策依据。

二、考试评价与考试质量分析的比较分析

(一) 对考试的界定不同

考试评价将考试界定为考试本身是人们希望拥有的价值的一种表现形式,而考试质量分析则将考试界定为测量是否具有某种既定特质内在结构的活动。按照考试评价的界定,考试有无价值完全取决于考试自身能否与特定的对象建立价值关系,因此,考试实际上体现的是选用考试者需要的对象化,具有个性和相对性。但是,按照考试质量分析的界定,考试的质量与人的主观需要无关,完全由独立于人的考试(测量)规律决定,是考试合乎考试规律的表现,具有普遍性和绝对性(这里所说的绝对性是指考试作为一种测量工具具有客观存在的特性。当然,从考试技术自身发展的角度看,这种绝对性也具有相对的一面)。

(二) 活动的目的不同

前面的分析已经指出,考试评价的根本目的是针对特定的选用考试者,对特定考试的合乎考试目的性和合乎考试规律性进行判断,以便人们能合乎价值规律地科学设计和选择考试。考试质量分析的目的在于判断考试的信度、效度、难度、区分度等指数是否达到有效、可行之作为测量工具的考试的质量要求,发现可以进一步改进该考试的方向、策略和方法,以及找出达不到要求的原因和补救方案。

(三) 结果所依据的标准不同

考试评价和考试质量分析所依赖的标准不同。考试评价所依据的标准是根据选用考试者所意识到的需要及实际可操作性而建立起来的评价标准,这个标准具有相对性和不确定性,不具有普遍适用性。考试质量分析依据的标准是

考试所构造的特质合乎其所欲测量特质及能够测出这种特质的程度,这个标准具有绝对性、确定性和普遍适用性。

(四) 实施者的价值观影响结果的程度不同

在考试评价中,评价者的价值观对考试评价的结果会产生影响,特别是在考试评价的评价主体按照自己的要求建立评价标准实施评价时,将对考试评价的结果产生很大的影响,有可能得出完全偏离选用考试者之价值标准的评价结论。但是,考试质量分析是价值无涉的技术性工作,考试质量分析的实施者本身是技术分析人员,必须尊重进行这项工作的技术规范,否则会出现科学性错误,但绝不会出现价值判断方面的错误,因此,考试质量分析之实施者的价值观不影响考试质量分析的结果。

(五) 考试对象的动机不同

从考试对象参加考试的动机的角度看,在作为考试评价对象之考试中,考试对象参加考试的动机是获得考试所表明的价值,这样的考试有时具有高利害性,导致考试对象与考试的实施者之间进行博弈。但是,在作为考试质量分析对象之考试中,考试对象几乎等同于物理测量中的被测量对象,没有任何主观动机可言,更没有与考试的实施者之间进行博弈的必要,对于他们而言,考试是低利害性的(甚至是没有利害性的)。

三、考试评价的基本框架

从考试评价所包含的内容来看,它主要包括分析考试自身的科学性、分析考试的情境适应性、分析考试对选用考试者的价值、分析选用特定考试的代价四个方面的工作。限于本文的篇幅,这里只概略讨论各部分的内容。

(一) 分析考试自身的科学性

本部分主要以考试作为测量工具所必须具备的结构和满足的测量规律为依据,分析考试本身是否符合考试之作为测量工具的科学性要求,相当于从测

量学的角度进行考试质量分析,主要包括从以下几个问题角度来分析考试的科学性:

——有无明确的考试目标?有无检核考试目标的效度标准?

——达成考试目标的策略与途径是否明确?对策略与途径的有效性和可行性是否有严格的论证?

——所选择考试题目的效度与信度如何?整卷的效度与信度如何?

——给出考试结果的方法是否科学?表述考试结果的形式是否合适?

——被评价的考试是达成本目标同类考试中最先进的方法吗?

——被评价的考试还有哪些不足?有无改进必要?如有,该怎样改进?

(二) 分析考试的情境适应性

本部分主要分析被评价的考试所蕴含的考试目标所规定的问题情境与实际情境的一致性,目的是判明考试所适用的范围,避免滥用考试,主要从回答以下问题的角度进行分析:

——该考试对考试对象的条件要求是什么?实际的考试对象是否满足考试的要求?

——该考试对相关的测量材料和配套的测量工具的要求是什么?实际情境是否满足该考试的要求?

——所选用的具体情境中还包含除考试目标以外的目标吗?如果包含除考试目标以外的目标,那么这个目标影响达成既定考试目标的途径和方式是什么?有无根本性的影响(即导致低效度,甚至是零效度的影响)?

——有两个或两个以上的考试目标与问题情境所蕴含的目标对应吗?如果有,那么根据什么确定考试情境所提供的信息符合被评价考试的考试目标?

(三) 分析考试对选用考试者的价值

本部分主要根据选用考试者对考试的需要为价值标准,对特定考试满足其价值的状态进行分析,是在一对一(一个选用考试者和一个价值客体)情况下进行的评价活动,主要从回答以下问题的角度进行分析:

——考试目标是否符合选用考试者的目标需要?

- 考试程序是否体现选用考试者关于考试程序文明的要求？
- 考试结果是否体现选用考试者关于考试结果公正性的要求？
- 考试过程与结果的误差是否控制在选用考试者所允许的范围之内？
- 考试本身是否会导致有违考试所蕴含价值本意的考试对象追求价值的活动？

（四）分析所选用考试的代价

本部分主要根据选用考试者实际选用考试的标准及实际考试的效果，分析具体选用的考试满足选用考试者的价值情况，目的是判断所运用的考试是否达成既定的价值，具体实施考试的过程中是否产生对选用考试者有价值的额外收益，实现了选用考试者的价值最大化，以及考试中存在的问题和改进的策略，主要从回答下列问题的角度进行分析：

- 选用考试者具体采用的选用考试的标准是什么？为什么会有这样的遴选标准？
- 被选用的考试除了对选用考试者的价值之外，对社会整体是否有方法性或者其他贡献性的价值？
- 被选用的考试是否导致了意想不到的具有重要社会后果的问题，如导致了片面追求掌握知识形式而轻理解知识形式所蕴含思想、方法等简化学习过程的应试学习现象等？
- 被选用的考试的可操作性怎样？是否是在所有备选考试方案中，根据成本最低、效益最高、可操作性最强等指标综合评价所得出的最优方案？
- 被选用的考试的成本（含人力、物力、财力及其他社会资源、人造资源和自然资源）预算是否科学？
- 根据被选用考试已执行的实际情况，判断是否执行了既定的考试成本预算？判断该考试的成本预算是否科学及有无改进余地？
- 所选用的考试形式是否有利于选用考试者通过用考试来指导考试对象的决策和行为？

四、开展考试评价需要解决的若干问题

考试评价作为一个方法的系统,从推进评价工作的角度看必须解决好为谁评、谁来评、评什么、怎样评等四个具体问题。

(一) 为谁评

为谁评是考试评价需要解决的第一个问题。解决这个问题的本质就是确立考试选用者在考试评价中至高无上的地位,它是合理、合乎逻辑地开展考试评价活动的价值的、逻辑的前提。同时,它也为制定评价考试评价自身的评价标准提供了依据——考试评价必须有助于选用考试者把握自己所选用考试的价值。

(二) 谁来评

在考试评价活动中,由于考试评价自身是一项专业性很强的工作,尽管人们尊重选用考试者在考试评价中的地位与作用,但是,并不是任意选用考试者均能把握其所选考试对自己的价值的具体情况,有的选用考试者要把握自己所选考试的价值,就必须借助于专业人员的帮助。这就涉及由谁来组织实施考试评价的问题,即确定考试评价的评价者(主体)问题。

解决谁来评问题的答案有二:其一,选用考试者自己设计、组织、实施考试并对考试进行评价,自己把握考试对自己的价值;其二,选用考试者委托他人或有关组织来替自己设计、组织、实施考试评价,把握考试对自己价值。

第一种答案所提供的评价者与选用考试者具有同一性,其结果只会受选用考试者的评价专业水平影响。第二种答案所提供的评价者与考试选用者之间存在交互作用,其结果既受到评价者的专业水平的影响,也受到评价者与考试选用者交互作用质量的影响,更受到评价者自身的价值倾向影响。

对于第二种答案情形,要防止考试评价异化为以考试评价之评价者为考试的价值主体的评价,关键是选用考试者在众多能提供评价服务的专业人员或组织中,选择合乎自己意愿的人员或组织作为考试评价的评价者,一方面要使得

自己能有效地监控评价者的评价活动,另一方面要使评价者能自由地展示自己的评价能力。为此,考试选用者应与评价考试者签订具有法律效力的委托评价协议书,该协议书要明确评价者开展评价活动的原则、拟采用的评价策略和方法、获得评价报酬的方法与数量、双方承担责任的范围与方式以及评价过程中所发现的协议内容外必须解决问题的处理方式等内容。自由身份的评价专业人员或组织,依据委托评价协议书依法获得特定考试之考试评价的评价者身份。

(三) 评什么

这个问题涉及依据什么确定考试评价的内容和以什么样的评价内容评价考试评价:以考试研究的新领域作为得出考试评价结论的依据。确定评价标准是解决这个问题的前提,建立与这个标准相联的考试之直接材料、现象或考试之间接材料、现象(这相当于建立评价指标体系)是确定评什么的关键。

由于不同选用考试者对特定考试的需要不同,实践中很难建立一个对所有选用某考试者均适用的评价该考试的评价标准。对任何考试评价而言,实际的评价标准均必须征得选用考试者的同意,只有该考试的选用者与评价者之间进行充分的交流与沟通,共同建立评价标准,才能确保实际采用的评价标准有效、合理、可行。

(四) 怎样评

这个问题包括:考试评价过程中选用考试者如何保证考试评价者体现自己的价值标准、考试评价者用什么方式方法收集和解释评价信息、考试评价者用什么方法和形式报告评价结论等主要问题。

在考试评价过程中,评价者只对选用考试者负责。选用考试者拥有对考试评价进程的知情权和咨询权,但在行使这两个权利时应该尊重评价者的评价专业水平,不得干预评价者按照委托评价协议书所规定的评价形式、评价范围和所提供的授权自主开展评价,不得授意评价者按照自己的需要做出虚假的评价结论。

随着评价活动的展开,评价者感到只有补充新的收集和解释评价信息的方

式、方法或报告评价结论的方法与形式才能完成评价时,必须先征得选用考试者的同意才能补充,不得擅自变更既定的收集和解释评价信息的方式、方法及报告评价结论的方法与形式。

张远增

内容提要

本书从评价学角度系统地研究了考试的构成要素与机理,并对教学型考试评价、测量型考试评价和评价型考试评价之考试的基本问题、考试评价的原则、考试评价的评价标准、考试评价的基本策略、考试评价的模式及考试评价的质量保障等问题进行了研究,形成了考试评价理论体系。在此基础上,本书从教育考试设计的程式、试卷结构设计、试卷量程设计、试测与调试及高中学业水平考试试卷设计等维度研究了教育考试设计,以及从问题概析、题型及其机制、试题及其机制、试题设计的内容和试题设计的基本策略等方面研究了试题设计。本书建立的考试评价理论体系,以及提出的设计教育考试和考试命题主张,对于科学设计考试、运用考试、评价考试具有基础性的意义。本书对于考试研究者、考试开发者、考试使用者、接受考试者及关注考试者均具有一定的参考价值。

目 录

第一章 考试概析/1

第一节 考试的起源/1

第二节 考试的演进/3

一、普通活动型考试/3

二、专门活动型考试/3

三、专业活动型考试/7

第三节 考试的构成要素/12

一、价值主体/12

二、考试主体/13

三、考试标准/13

四、考试客体/14

五、需要/14

第四节 考试要素的交互作用/16

一、二要素交互作用/16

二、三要素交互作用/20

三、四要素交互作用/25

第五节 考试的机理/27

一、考试的学习机理/28

二、考试的测量机理/36

三、考试的评价机理/47

第六节 考试的分类/48

一、单一标准分类/48

二、复合标准分类/50

第二章 教学型考试评价/51

第一节 教学型考试的基本问题/51

- 一、核心问题：实现有效教学/52
- 二、累积——为学习奠基问题/54
- 三、同化——学习内容主体化问题/56
- 四、顺应——学习者对象化问题/58

第二节 教学型考试评价的原则/60

第三节 教学型考试评价的评价标准/65

- 一、有利于激发学生进一步学习的兴趣/65
- 二、有利于学生了解自己的现状/67
- 三、有利于学生寻求改进和完善自己的学习方法/67
- 四、有利于教师提高自己的教学能力/68
- 五、有利于降低课堂教学成本/69
- 六、有利于提高教学效率和效益/70

第四节 教学型考试评价的基本策略/71

- 一、分析考试的内容与考法/71
- 二、分析题目难度设计的原则与方法/73

第五节 教学型考试评价的模式/75

- 一、评价模式的结构/75
- 二、教学绩效证据的构成/76
- 三、判断教学绩效证据的标准/78

第六节 教学型考试评价的质量保障/79

- 一、建立教学之价值主体资格制度/79
- 二、建立教学之评价主体资格制度/80
- 三、建立教学绩效证据清单制度/81
- 四、建立教学绩效证据管理系统/81