

# 试卷质量评估方案的改进研究

陈丽兰

(重庆交通大学人文学院,重庆 400074)

**摘要:**文章从高校二级学院教学管理人员的角度出发,针对现行的试卷质量监测评估存在评估方案单一、指标体系过细、评测方法过精、操作程序较繁问题,进行了全面的创新改进,构建了被动评估与主动评估、精确评估与模糊评估、点评估与区间评估相结合、相互渗透的试卷质量评估新方案。实践表明,该方案具有评估对象众多、评估指标简明、计算方法方便快捷、可操作性强等特点,评估结果可信度高。

**关键词:**试卷质量;试卷评估;优化设计;区间评估;统计分析

**中图分类号:**G642      **文献标志码:**A      **文章编号:**1005-2909(2012)02-0087-04

随着教学质量工程建设的不断深入,研究考试质量的管理理论、探索试卷质量的监控与评估已越来越明显地成为教学质量管理工作的重要任务的迫切任务。近几年来,不少教育管理工作对试卷质量的测评理论和方法进行了卓有成效的研究与探索。朱德全<sup>[1]</sup>指出,“教育评价是指以教育为对象,根据一定的目标,采用一切可行的评价技术和方法,对教育现象及其效果进行测定,分析目标实现程度,从而作出价值判断的过程”,并给出了教育评测的信度、效度、难度、区分度等一系列公式;李向东<sup>[2]</sup>给出了试卷平均成绩、及格率、标准差的计算公式与简单评测;宋青<sup>[3]</sup>给出了试卷考核质量评价体系;张芊<sup>[4]</sup>又提出了对最高分、最低分、平均分、标准差、频数分布、成绩分布、偏度、峰度等进行评估,并给出了正常、偏低、偏高的评测等级。

文章首先指出了现行的试卷质量监测与评估方案的不足,并对其进行了全面的创新改进,构建了一种被动评估与主动评估、精确评估与模糊评估、点评估与区间评估相结合、相互渗透,既科学合理又方便快捷的试卷监测与评估新方案。

## 一、试卷质量评估的现状剖析

为了提高教学质量,各个高校都十分重视期末考试的试卷质量,为此建立了一系列管理制度和相应的评估指标体系。通过评估,试卷质量有了明显提高,但也存在不少问题。

第一,少数命题教师由于缺乏教学经验或命题经验,试卷命题质量不高,甚至出现异常试卷。如试卷制作不规范、考试内容偏离教学大纲、考试重点不分明、题型与题量过少或过多、知识考点分布不均衡、试题过难或过易、及格率过高或过低等。

第二,部分阅卷教师由于责任心不强,阅卷质量不高。如:不严格按照评分细则评卷;试卷评分不规范,随意更改评分,或改分而不签字;登分、总分多加少算等。

第三,少数教学管理干部由于自身素质限制,对教育管理基本知识掌握不牢,对现代教育科学评测基本技术不熟练,导致对考试质量的监控力度不够。如:关注定性的宏观评估,而对定量的微观评估关注不够;对考试的最终结果关注较多,而对命题的质量把关不严;对阅卷的规范性、登分和总分的多加少算比较重视,而对平均成绩、及格率、成绩分布、难度系数、难度区分却重视不够;对考试各种信息的收集整理比较重视,而缺乏对各种数据的科学统计与分析,无法及时提供具有说服力的统计数据。

第四,自上而下的教学评估造成部分教育管理工作者和教师被动地接受评估,而主动、自觉有效的自评不足,使评估效果难于达到预期目标。

## 二、改进精确定量评估的基本原则

经过多年的教学评估,其指标体系已经比较完整,指标的设置比较科学合理,在实际的评估中收到了较好的效果。但是,由于评估指标公式繁多、计算复杂,对非教育管理专业的广大教师和教学管理人员来说难于掌握与运用,难于开展自评自查,影响了教学评估的效果。

科学的综合与简化是科学,评估方法的改进与创新也是科学。在简化评估过程中,应遵从如下基本原则:(1)必须保持评估目的不变、主要内容不变、主要观测点不变;(2)改进后的评估指标体系应具有完整性、科学性、权威性,具有较好的可靠性、可信性、稳定性;(3)构建的评估指标与评测方法应尽可能简单、方便、快捷,通过评估实践,能得到多数参与者的认可与欢迎;(4)按照改进后的评估方案所得的评估结论,应与按照精确定量指标评估结论基本吻合,无显著差异。

## 三、构建简易、方便、快捷的评估新方案

为了进一步提高试卷质量评估实效,有必要对现行的教学评估方案、指标体系、评测方法等进行科学的简化,构建一种既有一定的科学性与合理性,符合教学规律,又具有计算简单、容易掌握、操作简便的评估指标体系与评估方法。

第一,将单一的试卷质量评估改进为命题设计的自我评估、任课教师的自我评估、教学管理工作者的综合评估等多层次、多侧面、多视角的多重评估。试卷质量本身是一个多因素的系统工程,运用单一的评估指标体系和评估方法,很难做出科学的评判。

第二,把强化命题优化设计与评估作为试卷质量评估的关键点。试卷命题质量的好坏直接关系到学生成绩的评定、补考学生的认定、优秀学生的选拔

等重要问题,特别是量多面广的公共课、基础课试卷的命题,其命题质量的优劣,还涉及到二级学院甚至全校的考试质量问题。考试质量监控的第一关是命题质量关,这是全面提高教学质量最重要的一步,是减少命题的盲目性与随意性、提高命题质量关键的一步,是监控试卷质量、防止异常试卷最有效的一步。

第三,将定量指标的精确化评估改进为精确与模糊相结合的混合型评估。评估指标必须科学地量化。没有量化指标的评估,是粗放式的评估。如考试及格率的评估,个别教师认为及格率越高,表明自己的教学质量越好,及格率偏低,说明自己要求学生严格。这显然是不合理的,也是不科学的。及格率既不是越高越好,也不是越低越好,其评估必须科学地量化。但是,过于精确的量化评估有时会变成更加模糊的评估。

第四,将“点评估”改进为“弹性区间评估”。评判及格率的好坏,不是一个精确的“点”,而应该是公认的、合理的、具有一定弹性的“区间”。如及格率介于65%~85%之间,人们普遍认为是“好”,凡是偏离这个区间的,就认为是“不好”。再根据偏离的程度,将“不好”细分成若干等级。这种以区间作为评测范围的评估,称为区间评估。而区间的上下限,可根据不同课程、不同班级、不同教学要求作适当微调,称之为弹性区间评估。

第五,将成绩分布、方差、区分度、正态分布曲线的评估改进为按主要分数段分布的弹性区间评估。如将成绩分布的评估改为按[0,44],[45,59],[60,79],[80,89],[90,100]五个区间,统计其人数及其比例,然后作出评估。

## 四、试卷命题的优化设计

邹兆南<sup>[5]</sup>提出了关于试卷命题质量评估指标体系的优化设计的方案和方法。这是非常合理的,也是十分重要的。所有的考试课程都必须进行命题优化设计,设计的主要内容应包括:(1)根据学校关于考试的总体要求,制定该试卷组题的实施方案与细则。(2)根据教学大纲和考试大纲的要求,确定本次考试的重点与次重点,界定必定考、可考可不考、肯定不考的教学内容。(3)根据考生的实际情况,合理设计试卷的题型与题量。(4)根据教材各章的知识考点与教学时数安排,较均衡地进行分配,使考试分数值与之大体匹配。(5)设置平均成绩、及格率的期望值及弹性区间。如该试题平均成绩的期望值为70分,弹性区间为[60,80],及格率的期望值为75%,

弹性区间为[65,85]。(6)按照“知道、了解”“理解、掌握”“熟练掌握与综合应用”不同教学要求层次,设置相应题目分数值的比例。如该试题的(容易题):(较易题):(较难与难题)=2:5:3。(7)设置最低、平均、最高难度系数的期望值区间。如三个难度系数的期望值依次为[0.08,0.15],[0.25,0.35],[0.6,0.8]。(8)可设置少量(如<10%)有相当难度的题目,能把学习特别优秀的学生与广大学生区别开。

实行考试命题优化设计,这既是命题人进行具体命题的实施方案和行动准则,也是管理部门对命题质量实施跟踪监控与质量评测的重要依据。

### 五、试卷设计质量的评估

谭远顺<sup>[6]</sup>给出的评测试卷质量的数学模型是一个精确量化与模糊评测相结合的模型,非专业人员难于操作。对此笔者改进为简化的精确与模糊相结合的评估模式,对试卷进行评估。

考试结束以后,根据考试结果的相关统计数据,进行自评与他评,检验考试结果是否达到预期目标,是否与命题设计的期望值区间相吻合,并按照吻合程度给出评估等级。

#### (一)模糊指标的设置与评判

评估试卷质量模糊指标的主要观测点可设为:(1)试卷命题申报程序规范,手续齐全;(2)试卷制作规范,表述清晰、准确;(3)考试内容与教学大纲的吻合性好,没有偏题怪题;(4)题型题量适中,适合学生实际情况。

根据考试结果,完全达到指标要求的评为A级,否则,按照吻合程度的不同,分别评为B、C、D三个等级。

#### (二)量化指标的设置与评判

评估试卷质量的量化指标可设置为以下等级。

第一,及格率弹性区间评估:卷面成绩及格率为65%~85%者评为A级,55%~64%或86%~90%者评为B级,低于45%或高于98%者评为D级。

第二,成绩分布弹性区间评估:成绩为90~100分的占10%~25%,70~89分的占40%~60%,45分以下者>0%,若同时满足以上三条者评为A级;不满足其中一条、二条、三条者,依次评为B、C、D3个等级。

第三,平均难度系数弹性区间评估:平均难度系数为2.0~3.5者评为A级,1.5~1.9或3.6~4.0

者评为B级,低于1.0或高于5.0者评为D级。

第四,最高难度系数弹性区间评估:最高难度系数为6.0~8.5者评为A级,为5.5~5.9者评为B级,低于5或高于9的评为D级。

第五,最低难度弹性区间评估:最低难度系数为1.0~1.3者评为A级,0.7~0.9或1.4~1.8者评为B级,低于0.5或高于2.0者评为D级。

#### (三)综合评估

在上述9个评估指标中,若达到 $A \geq 7, D = 0$ 者,综合评分为A级;若 $A \geq 6, C \leq 2, D = 0$ ,综合评分为B级;若 $A \geq 5, D \leq 1$ ,综合评分为C级;若 $D \geq 2$ ,综合评分为D级。

### 六、任课教师对学生答题质量的自我评估

试卷质量评估的目的是提高教学质量,而提高教学质量的关键是任课教师。当试卷命题质量被评估为优良时,教师可根据考试结果,以教学班为单位,对自己所授班级的试卷进行自评,教师自评的指标体系一般可设置为六项:及格率(同前),平均成绩及分布(同前),各个大题(概念、理论、运算、应用)的得分比例,成绩在90分以上的人数比例、名单、典型实例,成绩在45分以下的人数比例、名单、典型实例,所授班级之间的成绩比较。

对以上指标进行模糊与定量相结合的自我评估,教师可总结教学经验,充分发挥教学特长,不断形成与升华教学风格,找出教与学两方面的差距和问题,拟定整改措施。这是提高教学质量的法宝,也是通向优秀教师的必由之路。

### 七、管理部门的综合评估与统计分析

#### (一)教学管理人员的监控职责

教学管理人员的职责充分体现在对考试质量的“监”与“控”两方面。所谓“监”,就是对本学院所有命题教师对命题质量的自评、任课教师对任课班级学生答卷情况的自评、阅卷教师对阅卷质量的自查、实施全方位的监察督促,并在此基础上对所有课程的考试质量综合评估。所谓“控”,就是对综合评估结果进行数据统计与科学分析。通过分析,帮助质量评估最好的教师总结经验,鼓励他们继续努力;对异常试卷和异常阅卷分析主观与客观原因,并责令限期整改。

#### (二)试卷质量评估的数据统计实例

从学校人文学院某学期期末考试的试卷中随机抽取5门课程,其统计数据如表1所示。

表1 人文学院期末考试5门课程统计

课程编号	样本容量/ 人	平均成绩/ 分	成绩分布				及格率	难度系数		
			90~100分	80~89分	60~69分	45分以下		平均难度	最低难度	最高难度
1	723	67.7	8.8	30.5	33.4	12.80	72.7	0.36	0.18	0.56
2	399	76.2	5.6	34.3	58.8	0.00	98.7	0.24	0.10	0.37
3	182	61.4	11.6	10.7	38.3	18.60	60.6	0.39	0.28	0.50
4	67	81.9	20.3	45.2	19.2	0.00	84.7	0.18	0.09	0.25
5	268	40.5	0.0	7.2	24.6	21.60	31.8	0.28	0.10	0.40

## (三) 试卷质量的综合评估实例

估,然后按综合评判标准,给出各门课程的综合评估

按照上述评估指标体系和评估方法,先分项评

等级,结果如表2所示。

表2 人文学院期末考试5门课程综合评估

课程编号	命题设计	试卷制作	平均成绩	成绩分布	及格率	平均难度	难度区分	阅卷质量
1	A	A	A	B	A	B	A	A
2	A	A	A	C	D	B	B	A
3	A	A	B	C	B	B	B	A
4	A	A	A	B	A	C	B	A
5	B	A	D	C	D	A	B	A

  

课程编号	试卷分析	试卷保管	评判标准				综合评 估等级
			优(A≥7,D=0)	良(A≥5,D=0)	中(A≥4,C+D≤3)	差(C+D≥4)	
1	A	A	A=8	B=2	C=0	D=0	优
2	B	A	A=5	B=3	C=1	D=1	中
3	A	A	A=5	B=4	C=1	D=0	良
4	A	A	A=7	B=2	C=1	D=0	优
5	C	A	A=4	B=2	C=2	D=2	差

经过领导和专家评议,上述课程1和4被评为优秀,建议给予表彰和奖励,对课程5评为差,建议限期整改。

## 八、结语

文章针对现行的试卷质量监测评估存在评估方案单一、指标体系过细、评测方法过精、操作程序较繁的问题,在文献[1-6]的基础上,进行了全面的创新改进,构建了试卷质量评估新方案。实践证明,改进后的试题质量评估指标体系与评估方法,既简单、方便、快捷,又不失科学性,评判结果符合教学规律,具有较好的可信度。

## 参考文献:

- [1] 朱德全,宋乃庆.现代教育统计与评测技术[M].重庆:西南师范大学出版社,1998.
- [2] 李向东,宋一中,李文生.试卷分析指标及其计算方法研究[J].中国高等教育评估,2004(2):56-58.
- [3] 宋青,蔡景一.试卷考核质量评价.中国高等医学教育,2006(11):30-31.
- [4] 张芊,汪利,丁国勇.基于教学管理信息平台的考试质量分析系统[J].高校教育管理,2007(1):73-76.
- [5] 邹兆南,刘艺.试卷命题质量评估指标体系的优化设计[J].内蒙古师范大学学报,2007(9):73-76.
- [6] 谭远顺,邹兆南.试卷质量的评测模型研究与实践[J].数学的实践与认识,2009(22):204-210.

## Research and practice on the improvement of evaluation plan for test paper quality

CHEN Li-lan

(School of Humanity, Chongqing Jiaotong University, Chongqing 400074, P. R. China)

**Abstract:** From the perspective of administrative staff in secondary college, according to the problems in the current test paper quality supervision and evaluation, such as limited evaluation plan, complex index system, over-meticulous evaluation method and complicated operation procedure, the paper carried out some innovation, built up a new evaluation plan with the passive and active evaluation, accurate and general evaluation, point and area evaluation. The practice shows that there are some advantages including sufficient evaluation target, simplified evaluation index, efficient and operable calculating method. The evaluation result is highly reliable.

**Keywords:** test paper quality; test paper evaluation; optimized design; area evaluation; statistics analysis