

高考语文阅读与写作合卷分时考查的实证研究

赵静宇¹ 佟威¹ 漆成明²

(1. 教育部考试中心, 北京 100084; 2. 华南师范大学, 广州 510631)

摘要: 语文写作试题是高考语文赋分最高的试题,但它的区分效果与所占分值并不相称。有教师和研究者认为,在目前高考语文阅读与写作合卷合时考查模式下,阅读试题挤占了写作试题的作答时间,如果改为合卷分时考查,将保证写作作答时间,从而可以提高写作试题的区分效果。据此设计测试,将合卷分时模式下的实验组和合卷合时模式下的对照组进行对比分析。结果显示,合卷分时模式下高考写作试题的平均分和标准差有所提升,但不如预设的效果明显。

关键词: 高考改革; 高考语文; 写作试题; 区分度

【中图分类号】G405

【文献标识码】A

【文章编号】1005-8427(2018)01-0014-6

DOI: 10.19360/j.cnki.11-3303/g4.2018.01.003

写作试题的历史可以追溯至唐高宗时期进士科加试帖经、杂文。写作试题诞生之初受到士子重视的原因在于:它的区分效果优于其他试题:“惟对策多可抄袭,帖经惟资记诵,别高下定优劣,以诗赋文律为最宜,故聪明才思亦奔凑于此。”^[1]新中国成立以来,高考语文写作试题一直是高考各科试卷中赋分最高的试题,但写作试题的区分效果与所占的分值并不相称。为了最大程度地发挥写作试题在高考选拔中的作用,正向引导高中写作教学,研究者提出了多种解决方案,阅读与写作合卷分时考查是其中之一。这种方法是否有效,不仅需要理论的推导,也需要实证的检验。

1 问题的提出

1.1 写作试题区分功能现状分析

与国外其他考试的写作试题采用等级制的评分方式不同,我国高考语文写作试题采用计分制,

通常为60分,占总分的40%。写作试题所占的分数比重与写作能力的重要性是匹配的。随着21世纪全球化、信息化的知识经济时代的到来,写作能力对一个人的全面发展、终身发展是必不可少的。无论是经合组织(OECD)构建的《成功生活和健全社会的核心素养指标框架》、欧盟(EU)发布的《终身学习核心素养:欧洲参考框架》,还是联合国教科文组织(UNESCO)发布的《全球学习领域框架》,都将母语表达能力列为重要的内容。写作试题所占的分值虽高,有效分值却偏低。60分的写作试题,分数呈正态分布时理论上最优的标准差应达到10分,实际的标准差却往往不超过5分。写作试题的标准差小,学生成绩的离散效果不理想,影响了写作试题在选拔过程中应起的作用,造成的直接后果就是高中不重视写作教学。许多高中一线教师和学生存在这样的心理:写作练不练分数差不多,练多练少分数差不多,不如省下时间去提高数学、英语成

本文系全国教育科学规划单位资助教育部规划课题“高考国家题库建设研究”(项目编号:FCB160610)阶段性成果。

【作者简介】赵静宇(1983—),女,教育部考试中心,助理研究员;

佟威(1984—),男,教育部考试中心;

漆成明(1992—),男,华南师范大学心理学院,在读硕士生。

绩。与之相反,高校教师普遍认为,学生论文写作存在较大的问题,有些甚至做不到基本的文通字顺,无法清晰地表达自己的观点;他们希望高考加大对写作能力的考查力度,以此推动高中的写作教学,提高学生的写作能力。高考作为连接高等教育和基础教育的桥梁,应发挥其选拔功能和反拨作用,切实提高写作试题的区分功能。

根据2014年9月国务院颁布的《关于深化考试招生制度改革的实施意见》和各省陆续公布的新高考方案,全国统一高考科目只剩语文、数学、外语3科。在这种形势下,只有提高高考语文的区分度,才能防止或缓解在后续招生录取过程中可能出现的学生分数拉不开差距的问题。高达60分的写作试题更是责无旁贷。

1.2 写作试题区分功能不佳原因探讨

影响写作试题区分功能的原因较为复杂,以下3种观点较为常见。一是高考语文测试的是母语表达能力,学生在母语表达能力方面并没有明显的高低之分,写作分数拉不开差距是正常的。假如这种观点成立,那么,高考语文作为选拔性考试,要做的不是尝试提高写作区分效果,而是适当降低写作分值,增加其他可以增加区分效果的考试内容的分值。因此,这种假设并不成立。因为从过程评价的结果看,日常教学中学生反映出来的写作水平差别非常大,从文质兼美到错字病句累累,各个层次均有相当数量的学生。二是学生写作水平差别较大,但阅卷打“保险分”的心理造成分数趋中,导致写作试题的区分效果不佳。这一问题较为复杂,本文不作讨论。

本文关注的是第3种观点。越来越多的一线教师和学者持有第3种观点。他们认为,学生实际写作水平差别较大,在考场上表现出的写作水平却大致趋同,造成两者之间存在较大差异的原因在于写作试题受到阅读试题的挤压,作答时间严重不足。解决这一问题的方法是将目前阅读与写作合卷分时考查的模式改为合卷分时考查,即阅读试题与写

作试题依然在同一张试卷上呈现,但作答时间分别加以限定。2014年1月,在教育部考试中心组织召开的高考改革研讨会上,多位心理测量专家、语文教育专家提及这一方案。在2014年两会期间,全国政协委员、华东师范大学历史系章义和教授就此专门撰写了政协提案^[2]。这里需要注意的是,高考语文除阅读试题和写作试题外,还有其他试题,如语言知识应用试题。对写作造成影响的主要是耗时较长、分值较高的阅读试题,因此,讨论阅读与写作合卷分时考查时,指的其实是非写作试题与写作试题的合卷分时考查。

从理论推导和经验判断的角度看,这种观点是正确的。在高考语文试卷中,写作试题是试卷的最后一道试题,必然受到前面试题,特别是文字量和思维量较大的阅读试题的影响。早在20世纪80年代,高考语文命题组专家已经认识到合卷合时考查模式下阅读与写作会产生矛盾:“现代文阅读能力的考查应兼顾速度问题,这是毫无疑问的,但在阅读、作文共属一张试卷的情况下,阅读材料篇幅过长,会相应地影响作文的时间。”^[3]当时的解决办法是缩减阅读材料的长度。在信息化时代的今天,二者矛盾再次凸显。一方面,无论是高校学习、社会工作,还是日常生活,都要求学生必须具备一定的阅读速度,能够在规定时间内快速地从繁杂的材料来源中提取出有效的信息,并经过判断、处理为己所用;另一方面,写作要求学生就相关问题进行深入思考,构思文章,再从形之于心到形之于文。学生没有充裕的时间展示自己真实的写作水平,来不及构思、匆忙下笔,呈现出来的作文千卷一面,相差无几。据此推理,将阅读试题与写作试题分开考查,应该可以达到限定阅读时长、考查阅读速度和保证写作时间充裕、让学生体现出真实的写作水平的双重目的。

从考试测量的角度来讲,写作能力是较为独立的一种能力,单独限定时间更有利于精准测量。国外成熟的语言测试类考试,如英国的A-Level、

IELTS,美国的SAT、ACT、GRE无一不是写作单独考查,并在成绩报告中提供单独的写作成绩。

合卷分时考查模式能否提高写作试题的区分功能?是否会带来其他问题?可否真正应用到高考改革的实践?为了回答上述问题,教育部考试中心于2017年在某省进行测试和问卷调查,并对数据进行了细致的梳理和分析。

2 研究设计

2.1 测试目的

进行测试的目的,一是了解合卷分时模式对学生作答时间分配策略的影响;二是了解合卷分时模式对高考语文总分、写作试题和非写作试题分数及区分效果的影响。

2.2 测试设计

为了获得学生作答时间信息,测试采用机考模式,分为实验组和对照组。其中,实验组采用合卷分时模式,非写作部分限时90分钟,写作部分限时60分钟,非写作部分提交成功后方可作答写作部分;对照组采用合卷合时模式,试卷限时150分钟,学生可自由分配答题时间。

为了保证测试的数据对高考具有借鉴作用,测试卷依据2017年高考语文大纲命制,试卷的结构、题型、题量与2017年正式使用的高考语文试卷大致相仿,写作试题采用常规的材料作文。围绕测试目的,设计了调查问卷,了解学生的作答感受。

评卷与高考正式评卷的流程、要求相同。写作试题除了采用传统的综合评分方式外,还采用了分项评分方式,以消除评卷误差产生的影响。综合评分的评分量表是多年来高考一直沿用的评分量表,分项评分的评分量表则是在综合评分的评分量表的基础上修订而成的,评价的维度和标准与综合评分一致。评卷教师不直接给分,而是根据各维度将作文划分出不同等级,由计算机后台按照一定的规则自动计算分数,以此降低打“保险分”的心理造成的评卷误差。

2.3 抽样方法

根据测试目的和设计,选取信息技术条件较好的6所高中学校作为测试的样本校,覆盖质量较好、中等和一般的学校。采取整班抽样的方法,每所高中学校选取至少100名高二年级或高三年级学生参加测试,其中30~50名学生参加分卷测试。参与测试的学生共832名。由于系统原因导致一部分学生数据不全,剔除掉这部分数据后,最终收集到629份有效数据,其中分卷测试即实验组168份,整卷测试即对照组461份。

3 统计数据分析

测试完成后,对收集到的分数、时间、问卷调查结果数据进行综合分析。

3.1 合卷分时模式对学生作答时间分配策略的影响

根据计算机后台记录的学生作答时间,无论是实验组,还是对照组,学生的平均作答时间都在125分钟以内,见表1。语文考试时间为150分钟,学生答完全部试题后,还有约25分钟的时间剩余。仅从表1判断,合卷分时模式不会对学生作答时间产生太大的影响。

表1 学生作答时间统计

组别	实验组	对照组
作答平均用时/秒	7 245.68	7 383.42

进一步对作答时间进行细分后,发现事实并非如此。学生的阅读速度、答题速度和写作速度存在较大的差异。合卷合时模式改为合卷分时模式后,虽然总体作答时间不变,但是学生在整张试卷中的时间分配策略出现显著的变化。将非写作部分与写作部分的作答时间分开,作出学生作答时间分布图(见图1和图2)。在图1中,在对非写作试题进行作答时,相较于对照组,实验组学生的时间分布上出现了明显的双峰,集中在85分钟和90分钟左右。可以推测,实验组部分学生因时间限制改变了考试策略,在系统预设的90分钟内规划作答;部分学生

因时间限制未能及时调整考试策略,90分钟时被迫中断作答。在图2中,在对写作试题进行作答时,相较对照组,实验组学生时间分布上有一个明显的断崖,出现在60分钟。可以推测,实验组部分学生因时间限制,在系统预设的60分钟时被迫中断作答。在对照组中,非写作部分超出90分钟和写作部分超出60分钟的学生都不在少数。但在合卷合时模式下,学生有选择优先作答哪一部分试题和根据自己情况安排作答时间的自由。在合卷分时模式下,这种自由将会部分丧失。

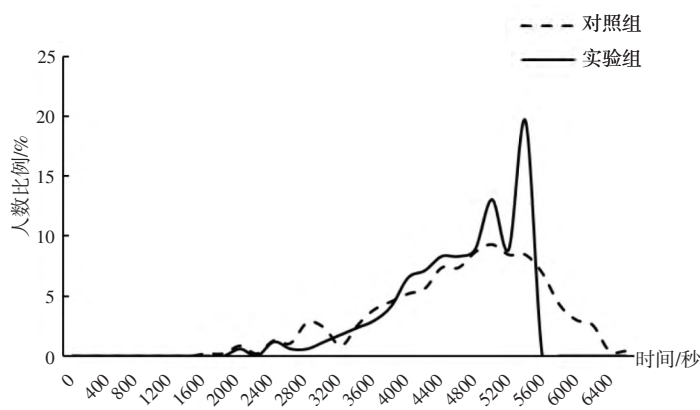


图1 学生非写作试题用时分布

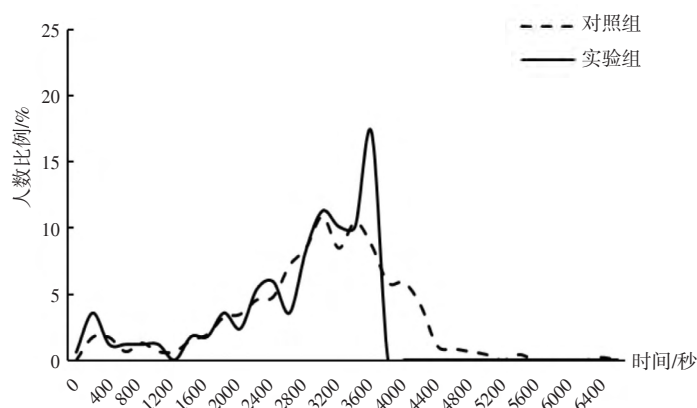


图2 学生写作试题用时分布

预设采用合卷分时模式的实验组学生普遍感到写作时间变得充裕的情况并没有出现。在对调查问卷中“是否有足够时间完成写作”的问题进行统计时(见图3),实验组选择“轻松完成”的学生仅

占29.90%,60.82%的学生“勉强完成”,9.28%的学生选择了“无法完成”,虽然优于对照组,但效果并不明显。

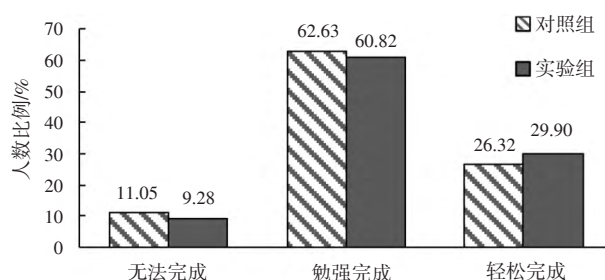


图3 “是否有足够时间完成写作”问题的调查结果

写作时间可以分为3种。第一种是考试规定的写作时间。在合卷合时模式下,考试规定的写作时间理论上可以长达150分钟;在合卷分时模式下,考试规定的写作时间是60分钟。第二种是学生预期需要的写作时间,即学生根据平时的练习预计自己将会使用的写作时间。第三种是学生实际使用的写作时间。在理想状态下,3种写作时间是一致的,但在现实中,常常存在较大差异。比如,在实验组,考试规定的写作时间是60分钟,学生出于留出检查时间等心理,将预期的写作时间定为50分钟,但实际使用的写作时间可能仅有45分钟。在这种情况下,虽然考试规定的时间尚未用完,但学生依然会有紧迫感。上述推测可以在调查问卷中另一个问题的数据中得到佐证。学生考试实际使用的写作时间与日常训练花费的时间之间存在较大的差异:实际考试用时明显低于日常训练用时(见图4)。

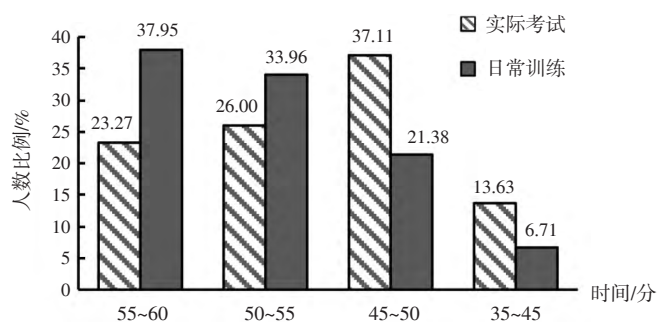


图4 学生写作试题实际用时与日常训练用时对比

习惯的养成非一日之功,时间分配策略的改变也不可能一蹴而就。在本次测试中,合卷分时模式改变的是考试规定的写作时间,学生实际考试用时随之调整,但学生预期的写作时间却没有改变。这是实验组学生即使拥有60分钟时间用于写作却依然感觉时间紧张的原因。

3.2 合卷分时模式对试卷分数及区分效果的影响

从测试的结果看(见表2),实验组的平均分和标准差都略高于对照组。写作试题部分未出现预期的平均分数和标准差都明显提高的情况。采用分项评分,实验组平均分比对照组高0.77分,标准差比对照组高0.41分;采用传统的综合评分,实验组平均分比对照组高0.51分,标准差比对照组高0.66分。有趣的是,相较实验组和对照组之间的差异,分项评分和综合评分之间的差异更为显著。在实验组,分项评分方式下的平均分比综合评分低3.37分,标准差高2.52分;在对照组,分项评分方式下的平均分比综合评分低3.63分,标准差高2.77分。

表2 试卷分数及离散程度统计

题型	实验组		对照组	
	平均分	标准差	平均分	标准差
写作试题(分项评分)	35.39	16.26	34.62	15.85
写作试题(综合评分)	38.76	13.74	38.25	13.08
非写作试题	47.14	10.39	46.20	10.27

在综合评分方式下,实验组和对照组写作试题分数分布的曲线几乎重合,特别是36分以上的部分(见图5)。

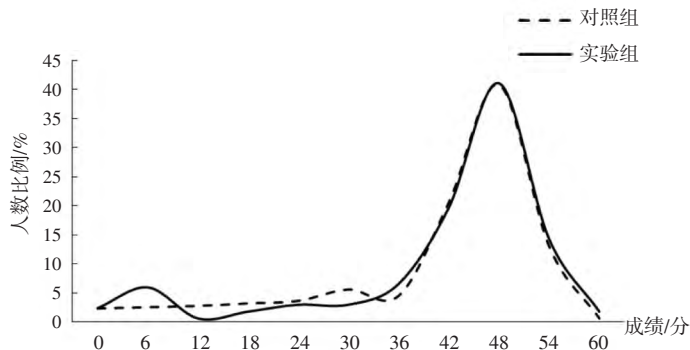


图5 写作试题分数分布(综合评分)

在分项评分方式下,虽然实验组和对照组的平均分相差无几,但分数分布却出现了较大差异(见图6)。首先是实验组的分数峰值变低,分数的整体分布更加均匀。在30分以下,对照组分数分布更加均匀,在30分以上,实验组分数分布更加均匀。在50分以上的高分段,实验组人数比例明显高于对照组。

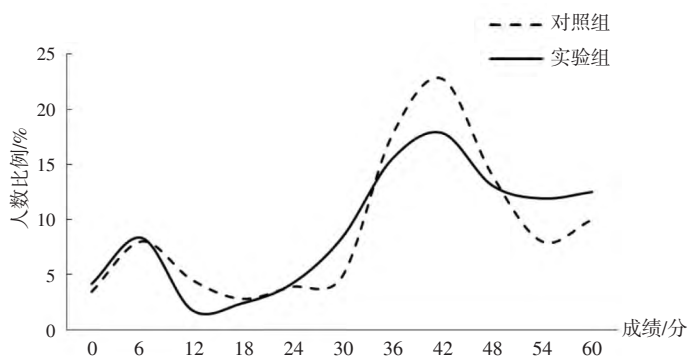


图6 写作试题分数分布(分项评分)

在非写作试题部分,实验组和对照组的分数分布均呈正态。相较对照组,实验组分数略有提高。但在72分以上的高分段,对照组人数比例明显高于实验组(见图7)。

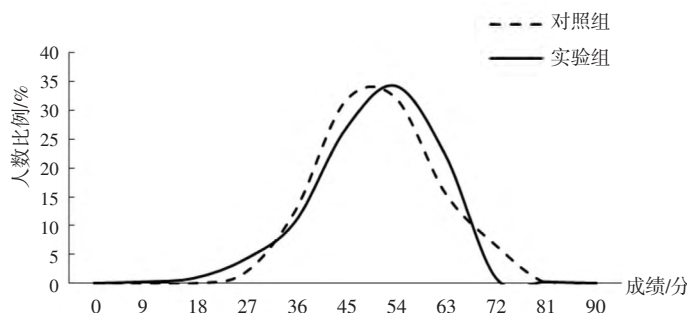


图7 非写作试题分数分布

4 结语

根据上述分析,可以得出以下结论。

第一,在分卷合时模式下,学生因模式更改调整了时间分配策略,但未明显感到写作时间的充裕。

第二,在分卷合时模式下,写作试题和非写作试题的平均分和标准差略有提高,但不明显。

第三,在分卷合时模式下,分数分布更加均匀,写作试题高分段人数增加,非写作试题高分段人数减少。

第四,评分方式改变对平均分和标准差的影响大于考试时间模式改变造成的影响。

在本次测试中,合卷分时模式产生的影响不如预期那样明显,主要是由于学生长期养成的时间分配策略较为稳定,不会因一次考试的改变而改变。要充分发现和利用合卷分时考查的预期优势,应在测试前即对实验组学生进行合卷分时模式的考试练习,使他们形成该模式下最适宜的考试时间分配策略,然后再比较实验组与对照组的异同。

合卷分时考查是一个涉及因素较多、较为复杂的实证问题。由于本次测试的样本量较小,测量获得的数据和统计分析结果仅可作为参考。在提高写作试题区分功能方面,需要尝试和探索更多的新方法。

参考文献

- [1] 钱穆. 国史大纲[M]. 北京: 北京商务印书馆, 1996: 430.
- [2] 人民政协报. 在沪全国政协委员“前瞻”全国政协十二届二次会议[N]. 人民政协报, 2014-03-02(1).
- [3] 高考命题组. 关于1986年高考各科命题的说明[J]. 人民教育, 1986(9): 24.

A Study on Assessing Chinese Reading and Writing in College Entrance Examinations in One Test Paper but in Separated Time Periods

ZHAO Jingyu¹, TONG Wei¹, QI Chengming²

(1. National Education Examinations Authority, Beijing 100084, China;

2. South China Normal University, Guangzhou 510631, China)

Abstract: In the college entrance examination, the writing item score is the highest, but its discrimination power is not good. Some teachers and scholars believe that reading occupies the time of writing, and if reading and writing in one test paper but in separated time periods, the discrimination power will be enhanced. According to this view, a test was designed. The experimental group which reading and writing in one test paper but in separated time periods and the control group which reading and writing in one test paper and the same time period were compared. The data show that average and standard deviation, the average and standard deviation of the experimental group have been improved but the effect is not obvious.

Keywords: college entrance examination reform; college entrance examination of Chinese; writing item; discrimination index

(责任编辑:周黎明)