

“教学评”一体化下考试评价模式优化研究

——以边民地区A县B中学高中英语学科为例

周钰婷

云南师范大学附属丘北中学

DOI:10.12238/mef.v7i12.9838

[摘要]“教-学-评”一体化教学理念下,学生是学习的中心,学生学习的过程是重点,教学评价是贯穿于教学活动全过程的“推进器”和“诊断器”。教、学、评三者间相互联系、相互促进。边民地区学生学习英语的时间相对较晚,学习兴趣不浓厚,英语基础较薄弱,而考试评价方式又较为机械、缺乏针对性,未能达到促进学生学习的目的。探索边民地区“教-学-评”一体化模式在优化高中英语考试评价中的实际应用,对提升高中英语教学质量和培养英语核心素养具有重要意义。本研究以云南省A县B中学为例,采用实证调查、典型个案访谈、文献研读、经验总结等方法,初步探索“教-学-评”一体化模式在优化边民地区高中英语考试评价模式的实施路径,并提出“中体量鼓励性评价”“中小体量针对性评价”以及“中大体量突破性评价”三个方法,以期对相关研究提供参考。

[关键词] “教-学-评”一体化; 边民地区; 高中英语; 考试评价

中图分类号: G633.41 文献标识码: A

Research on the Optimization of Examination Evaluation Mode under the “Integrated Strategy of Teaching, Learning and Evaluation”

—Taking the English Subject of Middle School B in County A in Border Minority areas as an Example

Yuting Zhou

Qiubei Middle School Attached to Yunnan Normal University

[Abstract] Under the “Integrated Strategy of Teaching, Learning and Evaluation”, students are the center of learning, the learning process of students is the focus, and teaching evaluation is the “propeller” and “diagnostic tool” that runs through the entire process of teaching activities. Teaching, learning, and evaluation are interrelated and mutually reinforcing. Students in border areas learn English relatively late, lack interest in learning it, and have a weak foundation in English, and the evaluation methods of exams are mechanical and in lack of specificity, which fails to achieve the goal of promoting students’ learning. Exploring the practical application of the “teaching learning evaluation” integrated model in optimizing high school English exam evaluation in border areas is of great significance for improving the quality of high school English teaching and cultivating English core ability. This study takes B Middle School in County A, Yunnan Province as an example, and uses empirical investigation, typical case interviews, literature review, experience summary and other methods to explore the implementation path of the “teaching, learning, evaluation” integrated model in optimizing the evaluation mode of high school English exams in border areas. Three methods are proposed: “evaluation for encouragement in medium volume”, “targeted evaluation in small and medium volume”, and “evaluation for breakthroughs in medium and large volume”, in order to provide reference for related research.

[Key words] integration of teaching, learning, evaluation; border minority areas; high school English; exam evaluation

前言

2020年10月,中共中央、国务院印发《深化新时代教育评价改革总体方案》,明确提出:“坚持科学有效,改进结果评价,强化过程评价,探索增值评价,健全综合评价,”为教学评价拓宽思路,引导方向。^[1]上世纪90年代中期,日本学者水越敏行以美国著名教育心理学家布鲁姆的教育目标分类学为基础,创新了教学评价思想,首次提出“教-学-评”一体化概念。他认为评价不仅仅是为教学目标服务,还应融入整个教学过程,以促进教师的教和学生的学。^[2]然而,对于“教-学-评”一体化的理解,有教师将其字面地理解为“教学+评价”,也有教师将其局限于某一具体教学环节或某一阶段的评价应用,还有把它简单地等同于各学年开展的大中小考试,等等。其实,教、学、评三个环节应连接形成一个闭环,有机贯穿于教学活动全过程、各环节,从而达到“以评促学、以评促教”的目的。^[3]

云南省文山壮族苗族自治州位于中国西南边陲的云南省东南部,它东与广西百色市接壤,南与越南社会主义共和国接界,西与红河哈尼族彝族自治州毗邻,北与曲靖市相连,是一个以壮族、苗族为主的多民族聚居地方,具有丰富的民风民俗和特色民族文化。A县B中学大力弘扬民族特色文化,积极吸收少数民族学生,现有32.2%壮族学生,28.6%彝族学生,14.3%苗族学生,12.1%回族学生,其余学生为汉族或其他少数民族。

下文将以A县B中学为例,指出边民地区(边疆民族地区)高中英语考试评价中存在的问题,以问题为导向,围绕“教-学-评”一体化理念在一学期、三阶段内考试评价的应用,提出几点针对性优化措施。

1 问题提出

A县B中学学生接触英语时间较短,67.1%的学生初中开始学习英语,29.4%学生小学阶段接触过英语,整体英语基础薄弱,应对纸笔考试能力较弱,对英语考试有很强的畏难情绪。据实地考察,该校主要以开展考试进行教学评价:高一高二年级每双周一周清考试,每月一月考,每季度一次的地区联考,再加上期中期末考;高三年级上学期每周一考,下学期“考一讲一”,试卷严格按照高考模式出题。该校高中三个年级整个学期的考试评价时长较长,四小门在40分钟以上,语、数、英在60分钟以上。在学期初阶段,学生答卷完成率较高,成绩基本维持上学期末水平;到学期中,学生空题情况越发明显,学生开始在考场上睡觉或随意填涂答案,成绩有了明显下降;到学期末,部分学生出勤率降到及完卷率明显降低,并有学生陆续办理退学,整体成绩与学期初相比进步甚微。

1.1 周清考试评价模式分析

据该校学部部长,该校周清考试每双周进行一次,考试题目由本年级教师自行出,试题来自“学科网”用户上传的习题。又根据该校高一年级周清考试安排表(见表一)和高二年级“周清”考试安排表(见表二),每双周学生要在星期六和星期日两天完成统考科目考试及试卷听讲,每场考试之间间隔10分钟,英语科目考试60分钟。教师则要求在这两天前完成出题和试卷印刷,

考试结束后进行试卷扫描、批阅及讲解,时间紧张,工作紧迫,难以对各个环节进行精细化处理。据该校英语教研组反映,教师再出题和选题时,基本没有时间先做一遍,没有时间把控质量和难度,题目及讲评备课是在监考时段抽空完成。

表 一

考试科目	星期	考试时间	考试科目	星期	考试时间
化学	星期六	14:20-15:00	数学	星期六	20:40-21:40
生物	星期六	15:10-15:50	物理	星期六	21:50-22:40
政治	星期六	16:00-16:40	英语	星期日	8:20-9:20
历史	星期六	16:50-17:30	地理	星期日	9:30-10:10
语文	星期六	19:30-20:30	试卷讲评	星期日	10:20往后

表 二

考试科目	考试时间		考试科目	考试时间	
语文	星期六	19:10-20:10	四选二	星期日	7:50-8:50
数学	星期六	20:20-21:20			8:50-9:50
英语	星期六	21:30-22:30	物理、历史	星期日	10:10-11:10
早读	星期日	7:20-7:40	试卷讲评	星期日	11:30往后

1.2 月考、期中考考试评价模式分析

A县B中学校把月考视作阶段性教学质量检测的重要指标,用于评价学生学习情况,进行划线排名计算上线率,也用于教师教学能力的评价和考核,与评优评先工作挂钩。据该校教务处安排,月考试卷由不同年级教师利用课余时间抽空按照高考题型出题,题目同样来自“学科网”用户上传的习题。其题目质量参差不齐,答案设置随意,重点难点与高考难以对接。以下是该校月考安排表(见表三)。

表 三

日期	时间	科目	日期	时间	科目
2023年12月28日	8:00—9:30	政治	2023年12月29日	8:00—9:30	物理
	10:30—12:00	生物		10:30—12:00	历史
	14:30—17:00	语文		14:30—16:30	英语
	19:30—21:00	化学		19:30—21:30	数学
	21:30—23:00	地理		8:00—9:30	物理

那么对于边疆民族地区高中阶段教育者而言,如何根据学生学习状态的变化,设计一个能鼓励学生学习、能进行针对性联系、能拔高能力的考试评价模式,形成教-学-评的学习闭环,就是本研究着重解决的问题。

2 研究设计

2.1 研究对象

研究选取A县B中学高中段学生为对象进行研究。该校高中共三个年级,每年级各10个班级,1至7班为物理方向,8至10班为历史方向,其中1、2、4、6、10班为重点班,3、5、7、8、9为平行班。(其中高一年级人数最多,共419人,样本数较多;高二年级人数减少,共397人,样本数中等;高三年级人数最少,共296人,样本数较少。)

2.2 研究方法

为探索适合边疆民族地区一学期内不同阶段应使用的考试评价方法,研究员采用分层随机抽样方法,分别在A县B中学高中段2023-2024学年春季学期三个阶段(学期初、学期中、学期末)开展追踪研究。

基于此,本研究共考虑年级、班级和分班方向三个层次,分别在A县B中学高一到高三各年级中抽取4个班(其中涵盖物理方向重点班1个、物理方向平行班1个、历史方向重点班1个、历史方向平行班1个),然后分发“阶段性检测知识问卷”,学生答题完成后统一回收并汇总进电子表册。最后,研究员分别在一、高二、高三各年级的“三、五、七月阶段性检测知识答卷成绩汇总表”中调取被抽取的4个班的成绩作为样本进行分析研究。

2.3 研究工具

本研究的研究工具为《春季学期三月份高中英语阶段性检测知识答卷》《春季学期五月份高中英语阶段性检测知识答卷》《春季学期七月份高中英语阶段性检测知识答卷》《三、五、七月阶段性检测知识答卷成绩汇总表》和SPSS16.0软件。知识问卷共包括六个不同体量的试卷,分别为“A卷(0-20分钟)”“B卷(21-40分钟)”“C卷(41-60分钟)”“D卷(61-80分钟)”“E卷(81-100分钟)”“F卷(101-120分钟)”,六种试卷类型的难度均与学生入学分班考试一致,题量从A往F逐渐增大。成绩汇总表包括“班级”“姓名”“入学分班成绩”“阶段性检测知识答卷成绩”和“进退分差”五项。

为客观地反映学生学习进退情况,“成绩汇总表”对比阶段性检测成绩和入学分班成绩的差距,采用六等级量表(Likertscale)进行量化比较:进步6分及以上为“明显进步”、进步6分以下、3分及以上为“轻微进步”、进退分差在3分以内为“没有进步”、退步6分以下、3分及以上为“轻微退步”、退步6分及以上、15分以下为“明显退步”,进步15分以上为大幅进步。

又由于各班人数不同,研究员在所抽取班级平均分发此6种知识答卷,每种共发出70份,每次样本总量为420个,三个阶段总共发出1260份知识问卷。

3 研究分析

3.1 学期初阶段

3.1.1 学情分析

学期伊始,同学们的心情是复杂而多变的,既有对新学期的

期待之情和兴奋之感,也有对未知挑战的紧张和焦虑。他们需要时间来调整自己的心态,适应新的环境和学习要求。对边疆民族地区学生而言,开学意味着不用干农活,他们对开学是充满期待的。以A县B中学为例,每班40余人中,仅有3到10名同学居住在县城内,依靠家庭成员外出打工取得工资性收入,其余学生家庭收入来源均是来自农业、副业。因此,学生在放假后一直到开学初期的一段时间内,大部分时间都是在田间地头干农事,制作农副产品,抑或从事副产品的售卖工作,以此维持日常开销和学业。

3.1.2 答卷成绩分析

(1) 信度分析。研究员先是对420名学生“三月阶段性检测知识答卷成绩汇总表”进行信度检测。该答卷工具由2个项目构成,分别为“试卷体量”和“进退幅度”。分析显示,Cronbach α 系数为0.788,系数接近0.8,^[4]表明项目之间存在着较高的内部一致性。这意味着所测量的项目在在一定程度上具有相似的特质,可以作为一个整体来进行评估和分析。

(2) 方差分析。根据方差分析结果显示(见表四),不同试卷体量组别之间的进退幅度均值存在显著差异, $F(5, 414)=256.778$, $p<0.001$ 。这表明试卷体量对考生进退幅度有显著影响。进一步观察各组间的均值差异,我们发现在0-20分钟组的考生进退幅度均值为1.10,而在41-60分钟组和61-80分钟组的考生进退幅度均值分别为4.74和4.11,且均为正向变化。

表 四

选项	样本量	均值	标准差
0-20分钟	70	1.1	0.3
21-40分钟	70	2	0.51
41-60分钟	70	4.74	0.85
61-80分钟	70	4.11	1.36
81-100分钟	70	3.9	0.3
101-120分钟	70	3.14	0.35
F	256.778		

(3) 结论分析。在实际教学实践中,教师应合理把控试卷体量,在最初阶段使用41-60分钟和61-80分钟体量的试卷,来让学生取得更明显的阶段性进步,以便增强学习信心。

3.1.3 建议应用考试评价模式: 中体量鼓励性评价

对于边疆民族地区高中学生而言,在开学前的一段时期,学生生活工作较为艰苦。部分壮族、苗族女学生还会受当地风俗习惯的影响,初中到高一阶段就开始生儿育女,并且假期期间的家务劳作也颇为繁杂艰辛。在此阶段,他们渴望上学,求学行为较积极,学习态度较端正,能以饱满的精神面貌和顽强的意志品质去面对学习上的困难。因此,在学期初,应使用中体量试卷,对学

生进行鼓励性评价让学生取得更明显的阶段性进步,从而增强其学习信心,坚定继续学下去的决心。

3.2 学期中段

3.2.1 学情分析

到了学期中,学生心情开始浮躁,学业压力不断增大,学校生活的规律性和高中知识的难度都让边疆民族地区的孩子颇为厌倦。在一次与学生代表交流的座谈会上,笔者了解到部分学生在学期中觉得自己很难取得进步,难以坚持学到放假,而且家里的农活也需要有人来分担,于是就选择退学。在A县B中学,学期中到学期末阶段,特别是小假期过后,办理退学或转学中职院校的学生最多。据该中学一名退学学生,他退学的原因是英语学不明白,以他的能力,考试题目根本做不完,空题太多,分数太低。

3.2.2 答卷成绩分析

(1) 信度分析。根据所提供的数据,本研究样本量为420人,项目数为2个,Cronbach. α 系数为0.751。本研究的问卷在内部一致性方面表现良好,可开展进一步研究。

(2) 方差分析。根据方差分析的结果,笔者发现了试卷体量对学生成绩的进退幅度有显著影响($F(5, 414)=468.291, p<0.001$)。具体来说,在不同试卷体量水平下,学生成绩的波动存在显著差异。进一步的事后检验表明,试卷体量为61-80分钟和101-120分钟的试卷在学生成绩波动上表现出最显著的差异,且成绩变化为负向变化。

表 五

选项	样本量	均值	标准差
0-20分钟	70	1.06	0.23
21-40分钟	70	1.33	0.47
41-60分钟	70	1.91	0.28
61-80分钟	70	4.61	1.01
81-100分钟	70	3.46	0.5
101-120分钟	70	3.07	0.26
F	468.291		

(4) 结论分析。结果表明,试卷体量对学生成绩的进退幅度有着显著影响。较长篇幅的试卷可能会导致学生成绩波动较大,而较短篇幅的试卷则相对稳定。因此在此阶段,最好使用0-20分钟、21-40分钟、41-60分钟的考试卷。

3.2.3 建议应用考试评价模式: 中小体量针对性评价

对于边疆民族地区高中学生而言,相较于其他学科,英语的难度是比较大的,尤其是到了期中,随着各科知识难度和题量的增加,学生在成绩方面较难提升,即使竭尽全力了还是做不完题目,这会大大打击学生自信心。因此在这个阶段应当减小试卷

体量,有针对性地考察课上学过的内容,让学生有充足的时间思考和作答,提高考试评价的意义和信度。

3.3 学期末阶段

3.3.1 学情分析

据A县B中学学生反映,到了学期末,虽然各学科要求掌握的知识累积到最大值,但整体的知识水平是有一定程度提升的,对即将到来的假期充满期待,自信心相对充足。又由于在初期和中期通过鼓励性和针对性的考试评价对学生的自信心和知识储备有一定的巩固和提高,因此到了期末,便有能力去尝试大体量的试卷,同时大体量试卷也为学生成绩进步提供空间。

3.3.2 答卷成绩分析

(1) 信度分析。在本次分析中,所测的Cronbach. α 系数为0.860,表明所设计的问卷具有较高的内部一致性,即问卷各项之间的相关性较强。这意味着参与者在回答问卷时,各个项目之间的测量结果是一致的,结果的稳定性和可靠性较高。

(2) 方差分析。通过一元方差分析(见表六),发现试卷体量对考生进退幅度有显著影响($F(5, 414)=253.200, p<0.001$)。根据各时间段试卷体量的均值数据,可以看出,考试时间在41-60分钟、61-80分钟和81-100分钟的试卷体量均值分别为4.61、4.49和5.54,且为正向变化。

表 六

选项	样本量	均值	标准差
0-20分钟	70	1.03	0.17
21-40分钟	70	1.4	1.21
41-60分钟	70	4.61	1.17
61-80分钟	70	4.49	1.21
81-100分钟	70	5.54	1.14
101-120分钟	70	3.01	0.12
F	253.2		

(3) 结论分析。随着试卷体量的增加,考生的进退幅度也呈现增加的趋势,试卷体量对考生进退幅度有显著影响,试卷体量较大的考试往往会导致考生进退幅度增加。因此,在此阶段,应使用40-100分钟或更大体量的试卷,以提高学生成绩,增强学习信心。

3.3.3 建议应用考试评价模式: 中大体量拔高性评价

根据维果斯基的最近发展区理论,教育者应为学生提供带有难度的内容,调动学生的积极性,发挥其潜能,超越其最近发展区而达到下一发展阶段水平。^[5]期末阶段是学生知识掌握相对较好的时候,应使用大体量考试试卷,从而为学生的进步提供空间,也为优等生拔高提升创造可能。如果在这个时候依然使用小体量的试卷,则不利于学生的进步。

4 结语

教学评价在教学过程中起着目标导向的作用,在“教—学—评”一体化引领下,评价的目的是为了促进学生的学和老师的教,是为了鼓励学生学习,为了让教师更有针对性地开展教学,在现时边境民族地区高中英语考试评价中,应根据实际情况,教育工作者应有意识地及时调整优化考试评价模式,做到精细化出题出卷,使评价更能促进学生的学习。

[参考文献]

[1]中共中央国务院印发.深化新时代教育评价改革总体方案.[EB/OL]https://www.gov.cn/gongbao/content/2020/content_554488.htm

[2]李伟.教学评一体化视域下生物学项目化学习的优化[J].教学与管理,2024,(25):50-53.

[3]王璐.基于主题意义的英语“教—学—评”一体化教学设计——以译林版英语教材四年级上册Unit 4 I can play basketball为例[J].小学教学参考,2024,(21):71-73.

[4]吴明隆.问卷统计分析实务——SPSS操作与应用[M].重庆:重庆大学出版社,2010.

[5]Vygotsky,L.S.Mind in society:The development of higher psychological processes.[M]Cambridge,MA:Harvard University Press,1978.

作者简介:

周钰婷(1993—),女,汉族,广西合浦人,硕士研究生,中小学一级教师,云南师范大学附属丘北中学,研究方向:中小幼课程与教学。