

《中国英语能力等级量表》视角下的“课赛融通”研究 ——以全国职业院校技能大赛口语赛项为例

李明 卫蕾 徐红梅

(陕西机电职业技术学院)

摘要: 该研究主要用于检测某高校以“课赛融通”为指导的英语教学改革与《中国英语能力等级量表》对接的有效性。研究采用修订版 Angoff 法, 通过收集该校学科英语测评中的口语测试与《中国英语能力等级量表》的对接例证, 从而得出临界分值, 并针对内部效度和外部效度进行讨论。研究结果表明: 口语测试及格线对应《中国英语能力等级量表》中“口语能力量表”的五级标准。该研究为学生深入了解学习效果提供了详细的反馈报告, 并且为通过《中国英语能力等级量表》评测口语大赛与英语教学的融通研究提供了参考。

关键词: 课赛融通; 口语测评; 《中国英语能力等级量表》; 修订版 Angoff 法

Abstract: This study aims to provide validity of the alignment between the oral English subtest instructed by unity of lectures and competitions and China's Standards of English Language Ability (CSE). Modified Angoff Method is adopted in this research to obtain the critical value with evidence of internal validity and external validity discussed in detail. The result is as follows: the oral English subtest is aligned to CSE-5. On the other hand, the detailed feedback can be provided as well for the test takers in this study. It also provides an experimental reference for future research on unity between the oral English competition and lectures instructed by CSE.

Key words: unity between competitions and lectures; oral English Evaluation; China's Standards of English Language Ability (CSE); Modified Angoff method

1. 引言

全国职业院校技能大赛英语口语赛项(以下简称英语口语赛项)被称为职业教育英语学科的高考, 越来越多的职业院校将口语赛项的考核内容作为教学改革的主要内容。但是“为赛而教”的表象使得教赛脱节, 难以真正实现“课赛融通”为指导的教学改革。《中国英语能力等级量表》(以下简称《量表》)为中国外语教学和测评提供了统一标准(刘建达, 2018)。“将测评与标准对接……有助于清晰阐释分值的意义(Tannenbaum & Wylie, 2008)”。国外在二十一世纪初, 就有了针对考试和标准(《欧洲语言共同参考框架》, 2008)对接的大量研究(Tannenbaum & Wylie, 2008; Eckes, 2017)。而国内关于标准和考试的对接研究较少, 且尚未出现关于《量表》和口语赛项的对接研究。

“课赛融通”旨在将教学知识点和大赛技能点相融合, 从而使职业技能与课堂教学融会贯通, 是 2021 年全国职教大会提出的“岗课赛证融通”策略之一。其中“课赛融通”主要将大赛中的训练、方法、评价、内涵等和教学逐一匹配融合, 通过学生在技能大赛中提高能力素质来拓展课堂教学的育人效果, 实现课程教学和技能竞赛的融合统一。“课赛融通”与《量表》对接, 有助于针对英语学科制定科学的测评要求, 合理命题有助于学生对测评详细分析。同时发挥《量表》指导作用, “促进教学与大赛紧密结合, 推进英语教、学、评、考协调发展(王守仁, 2018)”。

试题中心法的典型是 Angoff 法(Angoff, 1971), 而“使用修订版 Angoff 法时, 需在首轮后提供考试题目难度值, 相比原始 Angoff 法, 修订版 Angoff 法可降低认知压力, 提高专家判断一致性”(Angoff, 1971)。

基于上述分析, 该研究旨在探索《量表》中“口语能力量表”与“课赛融通”的对接, 为《量表》在“课赛融通”推进中的指导作用提供依据。

1. 本文的“赛”专指全国职业院校技能大赛英语口语赛项, “课”专指英语学科教学。

2. 研究设计

2.1 研究工具

该研究以某高校“课赛融通”教改项目为工具。该项目是该校非英语专业学生通识英语必修课程教改项目, 考核方式对接英语口语赛项。考核内容包括演讲、图表分析、情境交流、职场辩论四部分。其中演讲和图表分析为个人口头表达, 其余两项为现场面试方式。考生通过标准为各项达到及格。为保证考核公平性, 同批学生均进行两次考核, 另外, 在对不同考核中的题目估算时, 采用双参数模型以保证考生得分具有可比性。

本文主要探究情境交流与《量表》对接, 对接素材包括 2020 年和 2021 年英语口语赛项全国赛题和结果数据。

2.2 参与人员

共计 10 位老师组成小组参与测评, 这些老师都是该教改项目参与者, 也是一线教学实施者, 具有权威性和代表性。其中 1 名组员是全国赛评委, 5 名组员担任过省、市级赛事评委, 熟悉试题类型及考生群体。

2.3 研究流程

该研究有 4 个阶段:

阶段 1: 熟悉《量表》中口语量表的描述语。

阶段 2: 分析赛题及其与《量表》的匹配度, 确定对接级别。

阶段 3: 设定标准, 通过三轮核准认定所教学生等级。

阶段 4: 效度验证, 通过研判小组成员评分的内、外部一致性获取效度。

依据组员初步判断, 将《量表》三到七级作为对接的主要级别。在以往研究中, “多轮判断被证明有助于提高结果的准确性(Tannenbaum & Cho, 2014)”。为保证数据效果, 此次研究中共进行

三轮判断,第一、二轮与三、五、七级对接研判,第三轮根据前两轮结果与四、六级研判。

2.4 数据分析

针对口语赛项和《量表》级别对接问题,以修订版 Angoff 法临界分值第三轮判断结果判定。参照同一领域内相关研究,组员研判结果的内、外部一致性仍以 FACETS 的标准模型(Many-Facet Rasch Model, Linacre, 2013)为准。此外,“衡量判断结果是否前后一致的重要指标是 Infit 和 Outfit 值(Linacre 2009)”,但在遇到极端数值时 Outfit 值更易受影响,因此本研究选择 Infit 值。如果某成员判断一致性不够理想,则在计算临界分时去掉该成员数据。

3. 结果

修订版 Angoff 法对考生做出了一致研判,本研究将按照 FACETS 分析结果、修订版 Angoff 法临界分值的顺序进行展示。

3.1 FACETS 分析说明

在 FACETS 分析结果中明显看出,除标号为 10 的组员外,其余组员的 Measr 值均在(-1,1)范围内,说明 10 号组员判断最为苛刻,超出可评价范围,因此去除其数据。raters 表示组员严厉程度。除了 8 号组员 Infit 值(2.44)超过 2 以外,其余组员 Infit 值均在(0.5, 1.5)之间,充分说明组员内部一致性较好,因此最终核准时删除了 8 号组员数据。

3.2 基于修订版 Angoff 法临界分值

	三级	四级	五级	六级	七级
计算结果	11.64	13.98	19.57	22.39	29.27
标准差	1.87	2.49	1.16	1.29	1.93
临界分数	8	12	19	22	31
百分比	29	53	66	79	89

表 1. 基于修订版 Angoff 法临界分值

上表是基于修订版 Angoff 法计算出的临界分值。其数据是依据组员一致性结果,去掉 8、10 号数据后得出。由于赛教融通教改实践的测评和考生学分以及毕业相关,考虑到通过率的问题,本研究舍去小数部分,只保留整数部分进行研判¹。

依据上表,组员在针对四级评价时标准差最大(2.49),说明在此级别的评审一致性较低。而在评价五、六级时标准差较小(1.16; 1.29),说明此级别的评审一致性更适合。且五级临界分值为 19,符合《量表》五级标准中规定的“需在考试中至少做对 19 题(63%题目)”(刘建达, 2018)的要求。本研究采用试题及格值为 18(等值后),和五级临界值最为接近,因此组员一致认同修订版 Angoff 法可将考核及格线对接《量表》五级标准。

4. 研判验证

通过研究,组员把基于修订版 Angoff 法考核及格线对接《量表》五级标准,虽然其临界值一致性较高,但结果效度仍需验证。本研究采用内部效度和外部效度进行讨论。

4.1 内部效度验证

主要采用组员内部一致性作为依据。表 1 中四、五级标准差(1.16; 1.29)集中反映了组员研判的一致性,充分说明各组员在这两个级别的评审更一致。其次, FACETS 结果仅有两位成员一致性不足,也表明小组成员具有较好的一致性。

4.2 外部效度验证

主要采用真实有效的学生测评。曾有研究质疑修订版 Angoff 法,

如 Eckes 称“‘最低能力考生’是虚假概念,不符合实际情况”(Eckes, 2017)。但在本研究中,组员通过多次、真实评测实践,证实了修订版 Angoff 法产生的数据是真实可靠的。

5. 结论

赛教融通是职业教育改革的新方案,能否以大赛评价标准作为教学测评指导方案,是本研究主要目标。《量表》的公布为英语教学和评价提供了统一标准”(刘建达, 2018)。本研究通过修订版 Angoff 法将赛教融通为引领的口语考核标准成功对标至《量表》五级标准,并通过内、外部效度佐证,证明了以大赛为指导的赛教融通是能够为英语教改及测评提供可靠依据的。

本研究对学生分析学业得失同样具有重要意义。学生可以依据所得分数,参照对接的《量表》描述语,了解优势与不足,完善学习目标。

此外,本研究为口语测评标准与《量表》评价标准对接提供了参考,也为对接研究提供了参考。

当然,本研究评价来源的途径较为单一,其评价数据的样本量不足,且本次研究的效度验证方式不够全面,因此其结论的广泛性和推广度还需要进一步深入研究。

参考文献:

[1]Council of Europe. Relating Language Examinations to the Common European Framework of Reference for Languages: Learning, Teaching, Assessment. A Manual[M]. Strasbourg: Language Policy Division, Council of Europe, 2009.

[2]Eckes, T. Setting cut scores on an EFL placement test using the prototype group method: A receiver operating characteristic (ROC) analysis [J]. Language Testing, 2017(3): 383-411.

[3]Tannenbaum, R. J. & Y. Cho. Critical factors to consider in evaluating standard-setting studies to map language test scores to frameworks of language proficiency [J]. Language Assessment Quarterly, 2014(3): 233-249.

[4]何莲珍, 罗 蓝. 语言考试与语言能力量表对接研究: 路径与方法[J]. 外语教学, 2020(1): 29-33.

[5]刘建达. 中国英语能力等级量表与英语测评[J]. 中国考试, 2018(11): 1-6.

[6]根据考试目的和特殊性选择合适的取整方法非常重要,否则会造成严重的后果(Council of Europe 2009)

[7]王守仁. 中国英语能力等级量表在大学英语教学中的应用[J]. 外语教学, 2018(4): 1-4.

基金项目: 本文系教育部职业院校外语类专业教学指导委员会外语教育教学研究课题(项目编号: WYJZW-2020-1047) 结题成果。

作者简介: 李明, 陕西机电职业技术学院副教授, 硕士, 教育部职业院校外语类专业教学指导委员会委员研究方向: 跨文化交际能力测评。

卫蕾, 宝鸡文理学院外国语学院讲师, 硕士, 研究方向: 跨文化交际研究。

徐红梅, 陕西机电职业技术学院副教授, 硕士, 研究方向: 教学法。