

研究初中综合英语测评中的信度与效度

张哲源

马来亚大学

DOI:10.12238/er.v8i1.5720

摘要：本文研究了信度与效度在庆来学校初中综合英语教学测试与评价中的应用，重点分析了如何通过科学设计测试与评价方案来提高测试的准确性和公平性。以云南省弥勒市庆来学校八年级的学生为研究对象，本研究首先进行了学情分析，并根据学生的实际情况制定了听、说、读、写四项技能的教学目标。测评方案分为形成性评价与终结性评价，结合英语演讲比赛、词汇竞赛等多项活动，对学生的综合英语能力进行了全面考查。通过平行卷测试法，进一步提升了测评的信度，并分析了表面效度、内容效度和结构效度在测评中的应用。最终，本研究得出结论，信度和效度在英语教学测试与评价中的应用是相辅相成的，科学的测评方案设计有助于准确评估学生的实际能力，为今后的英语语言教学提供了有力依据。

关键词：庆来初中综合英语；英语教学测试与评价；信度；效度

中图分类号：G62 文献标识码：A

A Research on the Reliability and Validity of Comprehensive English Test and Assessment in Junior High School

Zheyuan Zhang

University of Malaya

Abstract: This paper investigated the application of the reliability and validity on comprehensive English test and assessment in Qinglai Junior High School and analyzed how to improve the accuracy and fairness of test by designing a plan of test and assessment scientifically. The research object is some students from eighth grade in Qinglai Junior High School in Mile City, Yunnan Province. Firstly, this study carried on the study situation analysis and make teaching objectives from listening, speaking, reading, writing with students' skills. Formative assessment and summative assessment in evaluation project examined the comprehensive English skills with English speech contest, vocabulary contest, and other teaching activities. Secondly, parallel test method improved the reliability of test and analyzed the application of face validity, content validity, and construct validity in comprehensive English test and assessment. Finally, the result of this study has shown that the reliability and validity are complementary in English teaching test and assessment. Designing a scientific project of English test and assessment will be helpful to evaluate the real ability of students and provide a strong basis for English language teaching in the future.

Keywords: Qinglai Junior High School; Comprehensive English; English teaching test and assessment; Reliability; Validity

引言

本研究探讨了信度和效度在初中综合英语课堂测评方案中应用，特别关注这些措施如何提高评价的准确性和公平性。综合英语测评包括听、说、读、写技能，是了解学生综合英语水平和确定具体改进措施的基础。本研究以云南省弥勒市庆来初中为研究背景，采用科学设计的评价模式和测评方案对该校八年级学生的英语能力进行评价。本研究旨在分析在实施形成性评价和终结性评价，以提高英语测试的信度和效度。本研究通过英语演讲、词汇竞赛、平行卷测试等活动，探讨有效的测评方案如何有助于准确衡量学生的语言技

能，为今后的英语教学实践提供启示和建议。

一、测评方案概述

（一）学情分析

该班级来自云南省红河州弥勒市庆来学校八年级二班，学生的平均年龄为15岁，大部分的学生来自于城市，其少部分的来自于农村，班上共有32名学生，他们的英语水平都在中等，其词汇量大约在500，多数初中生已经达到《义务教育英语课程标准（2011年版）》四级水平。大家热爱英语学习，勤奋好学，大部分学生的英语成绩都处于稳步提高的状态。学生的英语这一单科成绩由刚入学时的平均分68，

到八年级平均分达到79分，同时最高分可达98分，最低分59。由于过分重视阅读、写作，忽视了听力和口语的重要性，所以学生在这两个方面处于劣势。本次测评中涉及的相关健康、环保的词汇表达有（health, doctor, clean, river）都在之前的课堂中有所涉及，并能准确的朗读和翻译^[1]。

（二）教学目标

首先，从知识目标层面来说，要求学生掌握情态动词、反身代词、比较级、最高级、一般过去时的被动语态的用法；熟悉句型表达 could you please/would like to do 的用法与考点；掌握每个单元的固定搭配、名词、动词、动词短语的辨析；熟练使用一系列健康生活、保护环境话题的词汇、句型。

其次，从能力目标层面来说，听力上，要求听懂文章、段落的大意，抓住关键信息，掌握精听和泛听相结合的听力技能，提高听力技能；口语上，准确流利使用目标语言和所学的新知识，提高学生英语公共演讲和口语交际的能力；阅读上，在阅读过程中，灵活使用阅读策略，读懂文章，把握不同体裁文章的阅读技巧，把握作者的意图和情感态度；写作上，通过对一系列范文的赏析、模仿和练习，把握不同体裁、不同话题的写作技巧，进而提升英语综合能力。

最后，从情感目标层面来说，个人方面，新冠疫情下的健康生活，学会解决健康问题，应对紧急健康事件；社会方面，培养学生助人为乐、为他人着想、乐于奉献、热爱公益事业、热爱劳动、创建文明卫生城市；国家层面，人与自然和谐共生，保护环境，创立环保世界^[1]。

（三）测评内容

本次测评方案是针对庆来学校八年级下册初中生开展的一次综合英语期中考试，测评内容出自人教版八年级下册第一、二、三、七单元和九年级全一册第十三单元，此次测评总分100分，主客观题相结合，全方位考查学生听、说、读、写的技能，形成性评价和终结性评价结合，各占比50%^[1]。

形成性评价围绕“迎健康生活，创环保世界”这一主题开展摄影比赛、英语演讲比赛、词汇竞赛三个活动。终结性评价以环保和健康为主题，从上述考查单元中，采用闭卷考

的形式，在120分钟内完成一套满分为100分的试题，包括听力部分、语言知识运用、阅读理解、写作四个部分。多项选择题、完形填空题、论述题、信息转换题、正误判断题、补全句子题几种题型^[2]。

二、信度

（一）概念

信度是指考试分数的一致程度，可以检验测试结果是否准确，具有准确性、一致性、稳定性、可靠性、公平性。它是任何一个有效测评中必不可少的考试要素^[3]。

（二）提高信度的方法

在形成性评价中，由于活动具有极强的主观性，为了保证形成性评价测评活动的有效性和公平性，信度的提高必须有多个老师依靠一套准确细化的评分标准进行测评打分。

在终结性评价中为了提高测评的信度，必须确保主客观题均衡分布于试卷中，可以采用重复测试法、平行卷测试法、对半分析法三种方法来确保信度的提高^[4]。在这里我们以平行卷测试法为例，对其加以细化阐述和介绍。平行卷测试法即老师出两套试卷A卷和B卷，这两套试卷的出题范围、考点、难度、试题数量、施考时间、题型一致，只是题目不同，先抽A卷让学生来考，隔一段时间对A卷评讲分析后，又让学生来考B卷，最终对比AB卷的考分，测定此次测评的信度的高低^[5]。

（三）信度在本次测评中的运用

以本次八年级期中综合英语测评为例，来研究信度在本次测评方案中的运用。

首先，在形成性评价中，以英语演讲比赛为案例，分析信度在其中的运用。本次英语演讲比赛，分为初赛和决赛两场，公布了一套细化的评分标准，初赛和决赛的评委都必须依据这套评分标准来打分，初赛由各班自行组织，各班英语老师为评委，每班选出得分最高的两位同学参加年级上的决赛。决赛邀请了七年级、九年级的全体英语老师和各班抽取的一位学生大众评审对12位选手打分。这样做就可以保证了测评的公平性和信度。

英语演讲比赛评分标准		
评价项目	评价要点	得分情况
演讲内容(15分)	1、思想内容能紧紧围绕主题，观点正确、鲜明，见解独到，内容充实具体，生动感人（5分）	
	2、材料真实、典型、新颖，事迹感人、实例生动，反映客观事实，具有普遍意义，体现时代精神和活动主题（5分）	
	3、演讲稿结构严谨，构思巧妙，引人入胜（5分）	
语言表达（15分）	1、演讲者发音标准，吐字清晰，声音洪亮圆润。（5分）	
	2、演讲表达准确、流畅、自然。（5分）	
	3、语言处理得当，语速平稳，语气、语调、音量、节奏张弛，思想感情起伏变化，能熟练表达演讲内容。（5分）	

形象风度（5分）	演讲者精神饱满，姿态、动作、手势、表情运用得当，合理表达对演讲稿的理解。	
会场效果（5分）	演讲有感染力、吸引力和号召力，能较好地与听众感情互动，营造良好的演讲氛围，演讲时间控制在3分钟之内。	
总计		

其次，在终结性评价中采用了平行卷测试法、主客观题相结合的方法提高测评的信度。通过如下表格，我们可以发现两套试卷各分数段的人数变化幅度较小，考分的一致程度高，信度高。此外，在本次考试中，总共有67题，其中多项选择题出现了30题，这种题型从四个选项选择一个正确选项，评分客观，操作简单，机器阅卷批改方便，信度也得到了保证^[6]。

各分数段人数		
	A卷	B卷
100-90分	8	9
89-80分	7	7
79-70分	8	8
69-60分	8	8
59分及以下	1	0

三、效度

（一）概念

效度就是通过测评是否测出了想要测的。一般来说，可以将效度划分为表面效度、尺度关联效度、内容效度、结构效度。表面效度是所想测的技能和测试的形式是否相匹配，比如想测写作能力，就让学生写作文；测口语，让学生参与英语演讲比赛，这样就可以具备表面效度。内容效度是指教学大纲、教学内容、教学目标所涉及和考试内容的关联性，学的是否考了。尺度关联效度即体现在AB卷上。结构效度是又要有了相应的题型，就体现了结构效度，例如有了阅读题，就有了阅读结构效度^[7]。

（二）效度在本次测评中的运用

1. 表面效度

在本次测评中，为了测试学生的口语能力，教师开展了英语演讲比赛，要求学生在3分钟之内，以健康、环保为话题，进行英语脱稿演讲，这样做可以测试出学生的口语能力，看出学生的发音、表达流利性和准确性，是否能够抓住要点讲解。

为了测试学生的词汇量和所学的词汇的掌握程度，教师开展了词汇竞赛，本次词汇竞赛题型为词义选择、词性转换、根据词义补全单词。通过这次活动，教师可以测出学生对于词汇的拼写、用法、意义的掌握。

为了测试学生的写作能力，教师在终结性评价的试卷中出了两道书面表达题，学生任选一道完成写作，且写作的内容与教学情感目标相契合，从这个题目可以测出学生的写作能力，对于写作内容、组织、结构、逻辑的把握，所以以上测评活动的测试技能和测试形式匹配，具有表面效度^[8]。

2. 内容效度

结合以上对于教学目标的阐述，可抽取试卷中的部分试题为例进行分析^[9]。

—Could I _____ your English book?

—Certainly, but you can't _____ it to others.

A. borrow; lend B. lend; borrow

C. borrow; keep D. keep; borrow

例如这个题考查了学生对于 borrow, lend, keep 三个动词的词义辨析，也考查了情态动词的用法，两个考点都在知识目标中所涉及，且是要求学生重点掌握的。

On April 4th this year, people all over China _____ (ask) to join in three minutes' silence to pay respect to the victims of the pandemic.

以上这个题要求用所给词的正确形式填空，不仅考查了一般过去时的被动语态的用法，而且涉及到新冠肺炎疫情，情感目标也得以升华。

从以上两个例子中，我们可以得出，本次测评内容覆盖面广、教学目标和考试内容具有较强的关联性，所以具有了内容效度^[10]。

3. 结构效度

本次终结性评价的题型分为听力部分、语言知识与运用、阅读理解、写作四个部分，所以就有了听力、语法、阅读、写作的结构效度^[11]。

四、信度和效度的关系

首先，信度是效度的前提和坚实基础，效度是信度的最终目的和归宿，二者缺一不可，相辅相成^[12]。其次，在测评系统中，假如信度高，效度可能高，可能低；如果效度高，则信度一定高；但假如效度低，信度可能高，可能低；最后，要想得到效度，必须保证信度，但信度可靠，效度未必会好。所以信度是效度的必要非充分条件^[13]。

例如，一位老师想要测试学生对于假期作业的掌握情况，开学后出了一套试卷，都是假期作业中所涉及到的内容，结果所有学生都考了满分，说明这套试卷的信度很高，但是缺乏效度，因为老师通过这套试卷不能测出每一个学生具体的

对于知识的薄弱点。所以通过上述例子，我们就可以看出信度和效度的关系。

[参考文献]

- [1]英语教学法教程[M].高等教育出版社,王蔷主编,2000
- [2]普通初中英语课程标准[M].人民教育出版社,中华人民共和国教育部,2018
- [3]英语语言测试的信度与效度[J]张蕾.张家口职业技术学院学报.2008(01)
- [4]交际语言测试:理论模式与评估标准[J].薛荣.安徽工业大学学报(社会科学版).2003(05)
- [5]粉笔教师.英语学科知识与教育能力[M].北京:新华出版社,2019:240
- [6]英语语言测试方法概述[J].马彩霞.校园英语.2015(01)
- [7]浅论语言测试的信度[J].陈纪梁,陶伟.山东师大外国语学院学报.2000(04)
- [8]论语言测试的信度和效度[J].温俭,杨华.河北软件职业技术学院学报.2007(03)

[9]试论英语语言测试中的信度和效度[J].宋波.时代文学(下半月).2010(04)

[10]教育测量与评价方法[M].吉林大学出版社,张洪秀,2014

[11]英语测试中多项选择题的利弊探析[J].卢书娟.湖北经济学院学报(人文社会科学版).2011(06)

[12]2019年高考英语全国卷语法填空题内容效度研究[J].刘超.教育测量与评价.2020(03)

[13]交际语言测试理论视角下英语专业四级中听力测试的结构效度研究[D].向前珍.西南交通大学.2014

[14]2008年西北地区初中毕业学业考试英语试题评析[J]崔秀玲.中小学外语教学(中学篇).2009(04)

[15]简明英语测试教程[M].高等教育出版社,邹申,2005

作者简介:

张哲源(2000.10.21-),男,汉族,云南弥勒,硕士研究生,职称无,研究方向:英语语言教学法,英语测试与评价,英语教学活动设计,英语语言学