

## 在语言测试领域运用 CEFR 的扎根理论方法论应用研究

高伟斌

湖南师范大学教育科学学院, 长沙 410081, 湖南, 中国

**摘要:** 文献调研表明, 扎根理论方法论 (GTM) 在语言测试领域鲜有应用。然而, 本研究在一项针对斐济一年级本科生写作能力的调查中运用了扎根理论方法论。该研究以《欧洲语言共同参考框架》(CEFR) 为标准, 成功对斐济某大学学生的写作能力进行了评估。CEFR 是全球最为完备的语言评估框架之一, 已被西方国家的众多语言测试机构广泛采用。本文并非着眼于研究所得的实证数据, 而是立足于方法论立场, 旨在引导扎根理论方法论的初学者将其应用于语言学领域, 尤其是语言测试研究。本文提供了分步操作指南, 并对佐证扎根理论方法论在语言学中之应用的相关文献进行了讨论。本研究的方法论贡献及其独特的数据集, 将推动学界与社会政策层面围绕这一议题的讨论。本研究就扎根理论研究方法论如何应用于纵向语言测试研究情境这一问题, 对知识体系作出了原创性贡献。目前, 高等教育中的语言测试主要依赖常规形成性评估与常模性评估所得的数据, 扎根理论与纵向研究设计等取向鲜有采用。本文对硕士与博士研究生等新手研究者以及政策制定者将该方法论应用于自身研究具有借鉴意义。

**关键词:** 扎根理论; CEFR; 语言测试; 备忘录; 学术英语

## Grounded Theory Methodology Using CEFR in the Field of Language Testing

WeiBin Gao

School of Educational Sciences, Hunan Normal University, Changsha 410081, Hunan, China

**Abstract:** The use of grounded theory methodology (GTM) is rare in the field of language testing as indicated by the literature survey. However, this study used GTM in a study on first year undergraduate students writing skills in Fiji. The study used the Common European Framework for Reference (CEFR) to evaluate student writing skills successfully at a university in Fiji. The CEFR is one of the most comprehensive frameworks for language evaluation that has been widely used by language testing organisations mostly in Western countries. This paper takes a methodological position rather than looking at empirical data from the research to engage novice users of GTM to apply it to the field of linguistics, specifically to language testing research. It provides step-by-step guide and a discussion of relevant literature that attest to the use of GTM in linguistics. The methodological contributions and the unique data set of the study will advance scholarly and social policy conversations on this topic. The study makes an original contribution to the body of knowledge on how grounded theory research methodologies can be applied to a longitudinal language testing research context. At present, language testing in higher education relies on data from conventional formative and

normative assessments. Approaches such as grounded theory and longitudinal research design have rarely been used. The paper will benefit novice researchers such as Masters and PhD scholars as well as policy makers in applying the methodology to their studies.

**Keywords:** Grounded theory; CEFR; Language testing; Memos; Academic English

## 一、引言

扎根理论（GT）是一种质性的、归纳式的研究取向，旨在探究、分析并生成关于个体行为、集体行动与社会过程的概念 [1]。它既是一种研究取向，也是一套通过数据分析建构理论的程序 [2]。扎根理论研究从一个宽泛的研究领域出发，允许理论从数据中自然涌现 [3]。早期具有开创意义的纵向研究，虽采用了访谈、观察与测评，但均未运用扎根理论取向。在此背景下，Rogers 支持在写作研究中开展纵向研究的主张，认为其对跨时间、跨情境之变化的关注，已被证明是理解写作发展的一种尤为适切的方法。简言之，在斐济高等教育部门的情境中，以扎根理论为视角的纵向研究一直付之阙如 [4]。

目前，大部分语言测试研究，尚未发现能够解释大学生写作水平差异这一现象的理论。Flick 指出，作为一种研究取向，扎根理论最适用于那些存在问题却缺乏解释的领域，也最适用于此前研究与理论建构相对匮乏、从而为新见解与新视角的生成留有空间的领域。Charmaz 对此表示认同，并补充说扎根理论为诠释复杂的社会现象提供了一种灵活而实用的取向。与其他方法论相比，扎根理论在质性研究中的独特作用在于其对理论建构的强调 [5]。采用扎根理论的另一个理由在于其对数据分析所采取的系统化取向。

Myers 指出，其他质性研究方法往往依赖宽泛的原则而非系统化的取向，这妨碍了它们的应用与诠释。因此，扎根理论的采用为本研究提供了创建新范畴与新概念的全新视角。由此，本研究对纵向研究与扎根理论研究的文献作出了新的贡献，填补了语言测试情境中的一项显著空白 [6]。事实上，可以说语言测试领域缺乏扎根理论是一项关键性缺口，因为它使教师缺失了实践与理论之间的纽带，并可能影响教与学过程的成效。

扎根理论研究在教育研究情境中同样颇为有用，因为它是在与人及实践的紧密结合中开展的，而这正是本研究的着力点。此外，Arthur 解释说，扎根理论有助于我们建构中层理论，此类理论在解释教育情境中的相关行为方面极具潜力，对教育场域中的教师与专业人员具有现实意义。笔者采用扎根理论研究设计，对本科生在大学一年级期间学术英语写作能力的进步（或缺乏进步）情况进行了评估。此外，本研究还采用了《欧洲语言共同参考框架》（CEFR）的写作语言标准 [7]。

CEFR 是全球最为完备的语言评估框架之一，已被世界各地的语言测试机构所采用 [8]。剑桥 ESOL、国际法律英语证书（ILEC）、Asset Languages 以及雅思等语言测试均与 CEFR 框架相对接。教学大纲设计者或语言测试提供方倾向于使其考试设计与 CEFR 相对接，原因在于该框架的透明性与连贯性 [9]。CEFR 还被应用于多种欧洲语言。因此，尽管 CEFR 未必尽善尽美，但它很可能是当前最为完备的语言评估框架之一。此外，CEFR 框架在学术语言测试评估中的全球性采用，使其成为本研究的理想选择。

以“纵向研究”与“扎根理论”两个短语在数据库中检索，未能获得任何与语言测试相关的结果，且在第二语言研究中亦显著稀缺。这印证了本研究采用扎根理论方法论的合理性。本研究所提供的见解，或可适用于其他以英语为第二语言的多语情境，为这些地区的大学制定教学语言政策提供参考，使其能够充分培养学生，使之在离开大学之后仍具备适当水平的书面语言能力。

## 二、《欧洲语言共同参考框架》（CEFR）

在语言测试领域，最具主导地位的框架之一便是《欧洲语言共同参考框架》。有证据表明，雅思等剑桥系

列考试将 CEFR 作为一项重要特征加以包含与体现。它们把 CEFR 纳入自身结构,并以多种方式加以呈现。在欧洲,CEFR 被用于服务促进语言多样性、资格透明化、劳动力流动以及终身语言学习等政策议程。在欧洲之外,CEFR 也已被广泛用于描述语言能力水平,并对当地的教学法与评估产生了相应影响。这使其成为本研究的理想选择,因为它为学生语言测试提供了能力水平的划分依据。

CEFR 还因其对质量保障事务的贡献而受到肯定,这种贡献不仅在于改进制度与程序,还在于支持语言学习、教学与评估相关人员及机构日益提升的专业化水平在论及质量保障事务时指出,CEFR 的《实践准则》为从业者群体提供了共同的参照框架与共享的元语言,用以反思和评估政策实践——确保改进之门始终敞开。

然而,对 CEFR 的应用亦不乏质疑。在语言测试界,对于把 CEFR 用作政策与实践协调统一之工具的做法尚存保留意见。批评者质疑 CEFR 在多大程度上为可操作的测试开发提供了适宜工具。例如,Nagai 等指出,马来西亚教育计划要实施 CEFR,需要拥有一支英语能力完全达标的教师队伍,并认为若教师本身英语能力不足,CEFR 便无法落实。马来西亚多数教师至今仍未意识到该框架,或对学习与采用该框架缺乏兴趣。CEFR 的作者之一 Coste 指出,通用等级描述语(如“B2 级学习者”)具有更高的便利性,而聚焦于特定活动的量表则更为精确却泛化能力有限(如某学习者在目标导向的合作中被判定为 B2 级,而在面向听众表达时为 B1 级),二者之间存在权衡。Coste 建议,为理解 CEFR 量表的局限,人们或许更宜将其视为“一种能够界定能力水平的测量工具,其刻度之精确犹如医用体温计上的分度”。Green 解释说,若无明确指引,“使用者可能会过度诠释 B2 级考试分数,把它视为在 CEFR 描述体系各个方面均具备相应能力的标志,而非仅仅是该测试实际所考查的有限能力子集”。

Green 提出的另一项论点涉及测试效度验证。Green 观察到,测试机构本应建构效度论证,以证明某测试用于特定目的的合理性(例如证明某潜在学生能够应对大学学习的语言要求),但如今存在这样一种风险:机构只需表明某项测试与 CEFR 相对接,便足以使使用者相信它几乎适用于任何目的。Fulcher 指出,CEFR 的批评者与辩护者均担忧分数使用者会把与该框架相对接之测试的结果视为可以互换,即认为“英国某测试的‘X’分与美国某测试的‘Y’分、欧盟某测试的‘Z’分在意义上相当”。Kolen 和 Brennan 澄清说,若要对两项或多项测试的分数进行等值处理(即把结果视为可互换),这些测试“在内容与统计特征上必须尽可能相似”。语言能力的多维性与争议性意味着,我们不能假定任意两项测试测量的是同一构念。针对上述批评与局限,CEFR 的作者们回应称,该框架的初衷是提供一种珍视并鼓励多样性的手段。

CEFR 作者之一约翰·特里姆教授在一次访谈中提到,在编制该框架时,其目标是形成一个共同的参照点,使从事不同领域工作、以完全不同的方式把它用于完全不同用途的人们,都能够以此为参照,从而感到自己属于同一体系。此外,针对 CEFR 缺乏建构测试细目所需详尽程度的批评,North 强调,该框架本身并非内容标准,而是一套生成性的“用以开发适合具体情境之差异化标准的装置”。欧洲以外采用 CEFR 框架的国家日益增多。Nagai 等报告称,越南、日本、印度尼西亚与马来西亚等国均已接纳 CEFR。在日本,大阪大学外国语学部开展了一项研究,以使该校所教授的 20 余种语言的课程体系趋于合理化。

在马来西亚,该框架于 2013 年正式引入,并被纳入《2013—2025 年马来西亚教育蓝图》与《2015—2025 年英语教育改革》之中,这表明政府不仅同意把该框架纳入并对接现行教育体系,还将加快其实施进程。除上述国家外,CEFR 亦对美国产生了吸引力。美国已设立了一种名为“Linguafolio”的《欧洲语言档案袋》(ELP)改良版本。斐济语言背景的特征与大多数以英语为第二语言的国家相似。在斐济,人们把英语作为第二语言使用,其母语则包括斐济语、印地语、罗图马语、日语、汉语或韩语等多种语言。这为本研究采用 CEFR 提供了空间,因为该框架具备在不同语言背景与语言需求之国家中加以运用的适应性。

### 三、相关文献

CEFR 的影响已延伸至语言教育之外的领域,例如把相关测试用于向移民授予永久居留权乃至公民身份。它被视为语言测试开发中的有效工具,并被公认为欧洲语言教育领域近期最为重大的事件。CEFR 已被用作众多语言测试研究的概念框架,例如 Byram 和 Parmenter 对法语、波兰语、保加利亚语、阿根廷西班牙语、汉语与台湾地区所用语言的研究,Diez-Bedmar 对英语的研究,Morrow 对法语的研究,Little 对爱尔兰语的研究,Fulcher 对欧洲诸语言的研究,Alderson 等对荷兰语的研究,以及 Hulstijn 等对荷兰语的研究。此外,Fulcher 认为,多年来语言测试已成为英国等欧洲许多地区的重要组成部分,CEFR 的影响力随之增强,因为它已被各国政府与教育机构用作语言测试的框架。

在语言测试研究中,《欧洲语言共同参考框架》自 2001 年问世以来,在界定能力水平(A1—入门级、A2—初级、B1—门槛级、B2—进阶级、C1—熟练运用级、C2—精通级)方面颇具影响力。《语言测试》期刊在其关于语言评估的特刊中,为 CEFR 在语言测试领域的影响力提供了有力证据。

Diez-Bedmar 在西班牙开展了一项研究,以考察同一院校群体内部的不同能力水平以及负向语言特征的性质。该研究运用 CEFR 能力水平与计算机辅助错误分析法,对大学入学考试英语部分的书面作文进行了检验。该研究共有 302 名被试参与,被要求围绕“在西班牙以外,你想去哪里作一次短途休闲旅行?”这一题目撰写作文。研究结果显示,多数中学毕业生在 CEFR 上的表现处于同一 B1 水平。该研究并未对各个能力水平的错误逐一分析,而是仅选取了 A2 水平,结果显示学生在情态助动词的使用上频繁出错。这为本研究提供了有益的框架,本纵向研究即运用 CEFR 能力水平对一年级与毕业年级本科生所处的水平进行了分类。

历史回顾部分包括对扎根理论方法论与纵向研究相关文献的细致梳理。扎根理论的一项重要优势在于,它直接植根于某一学科所面临的问题与议题之中,就本研究而言即语言测试。这一方法曾受到学者的批评,他们指出该取向中“扎根”“理论”与“发现”等概念,可能给研究者带来一种虚假的认识论安全感,同时削弱诠释、叙事与反思的重要性。这些学者认为,扎根理论强调秩序与程序性机制,这在很大程度上是其时代的产物,过去曾很好地服务于质性研究群体,但如今则可能被认为错失了质性探究中最精华的部分。尽管存在上述批评,扎根理论方法论所勾勒的方法论过程已被清晰证明是严谨而富有洞见的。Punch 和 Oancea 指出,研究者能够(逐步)看到分析性的故事线徐徐展开。基于此,他们强调扎根理论是目前应用最广、最受欢迎的质性研究方法之一。

### 四、扎根理论方法论的应用

Arthur 指出,扎根理论研究者会采用最契合研究问题的数据收集方法。扎根理论研究中常用的方法包括访谈、实地观察以及各类书面报告。数据收集分三个阶段进行:首先是开放式编码,其次是选择性编码,最后是理论编码。在开放式编码阶段,研究者贴近数据,保持开放心态去探索该组数据中正在呈现的模式。比较与提问是开放式编码中用以指导标签化的两项主要活动。随后,在选择性编码阶段,来自开放式编码的最重要编码被用作进一步数据编码的指引。最后,理论编码使研究者得以运用数据生成理论或一组主题,并加以整合。

扎根理论的最终成果并非一组研究发现或若干主题,而是一组围绕某一核心范畴/主题加以整合的扎根概念,由此形成一个理论框架,用以解释个人、组织、社区或国家为何以及如何经历并应对各种挑战或问题情境。换言之,它提供了一块基石,可据以建构知识与框架,用以指导实践。

### 五、研究阶段与数据收集

因此，本研究以 CEFR 为标尺，对一批一年级本科生在大学课程之初以及第一学年结束之时的学术英语书面能力水平进行了评估。第一次测试分为两部分：写作任务一要求学生对所提供统计表格中的信息进行概括，写作任务二则要求学生就给定题目提出论证理由。而在第二次测试中，第一项写作任务基于某公司的毛利润图表，第二项写作任务则要求学习者就该题目提出自己的观点并辅以相关例证。这使我们得以了解本科生在三年制大学课程第一学年中写作能力方面的差异与成就。由此，本研究的分析揭示了相关大学生是否走在以充分的学术英语语言能力完成学位的适当路径上。此外，语言测试的结果还可用于识别那些在书面语言技能方面需要额外支持的学生。

在这项为期一年的纵向研究中，共实施了三次写作干预。在第一次语言测试并依据 CEFR 完成评定之后，研究对象接受了旨在提升其写作能力的写作干预。为衡量写作干预是否奏效，第二次语言测试于其大学第一学年结束时实施。因此，该项目被划分为四个阶段，如下图 1 所示：



**Figure 1** The Four Stages of This Project  
**图 1** 本项目的四个阶段

在第一阶段，学生接受语言测试，其表现依据 CEFR 加以评定，以确定其语言能力与英语水平层次。写作干预随后构成项目的第二阶段。第二阶段的目标是通过各类写作任务提升学生的写作能力。在第三阶段，实施第二次语言测试并依据 CEFR 评定，以检验写作干预是否奏效。最后，第四阶段是对前三个阶段的总体评估，以提炼见解，这些见解构成了结论与建议的组成部分，可为解决本科阶段教学语言与知识可及性问题的相关政策提供参考，见图 2。

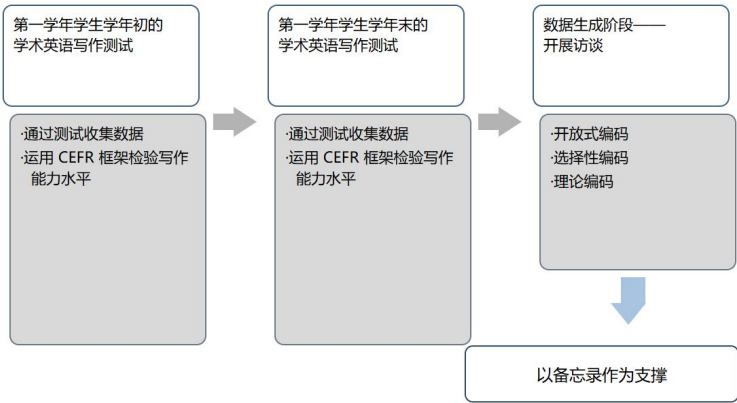


Figure 2 Data Collection Framework of Analysis Using Grounded Theory

图2 运用扎根理论的数据收集分析框架

第一学年结束时的测试完成之后，采用了相同的数据收集方法。此后，数据分三个阶段进行分析：开放式编码、选择性编码与理论编码。在开放式编码阶段，研究者贴近数据，保持开放心态去探索该组数据中正在呈现的模式。第二阶段是选择性编码，即把开放式编码中最重要的编码用作进一步数据编码的指引。最后是理论编码阶段，它使研究者得以运用数据生成理论或一组主题，并加以整合。

目前，斐济尚无关于大学毕业生为在就业与社会生活中有效运作所应达到之 CEFR 能力水平的统一规定。媒体上的招聘广告仅笼统写明“应具备英语能力”。雅思框架仅用于移民用途，并未纳入教育体系。

## 六、研究中备忘录的运用

在扎根理论方法论中，备忘录在整个研究过程中发挥着关键作用。本研究同样运用了备忘录。Birks 和 Mills 建议，研究一经构思便应开始使用备忘录。扎根理论学者 Charmaz 把备忘录写作视为启动并维持研究产出的一项技术。备忘录中所写内容既有通用性的，也有技术性的。可纳入备忘录的内容包括：

1. 你对自身研究的感受与假设；
2. 你在自身研究方面所持的哲学立场；
3. 与研究设计相关的潜在议题、问题与关切；
4. 对研究过程的反思，包括影响研究质量的各种因素；
5. 程序性与分析性决策；
6. 编码、范畴以及你正在建构的理论。

理论性备忘录不同于上文 Birks 和 Mills 所讨论的备忘录，因为它主要关涉实证数据。对数据中正在呈现之内容的讨论构成理论性备忘录的基础，在编码过程中尤其如此。Urquhart 建议：“无论你是否在从事理论建构式的研究设计，都应在编码数据时采用这一做法。”它是理论建构的重要工。在数据分析过程中，发现编码与范畴之间的关联对理论性备忘录颇有裨益。本研究采用 Birks 和 Mills 所提出的备忘录撰写指引来理解数据的通用性成分，并采用 Glaser 所提供的规则撰写理论性备忘录。学者们认为，备忘录的撰写并无标准格式，其形式灵活，取决于研究者本人。因此，笔者采用了一种简单格式，标明备忘录的主题，并区分其为通用性备忘录还是理论性备忘录，同时注明日期。

## 七、研究发现与讨论

如前所述，本文从方法论立场审视该研究，因此该研究的实际结果可参见 Goundar，本文的讨论仅限于扎根理论方法论与 CEFR 的运用。在学年之初，120 名学生中有 62 人处于 A1 水平、49 人处于 A2 水平，二者均归类为基础使用者。达到 A1 水平的学生性别比例几近均等，该水平记录到 27% 的女生与 25% 的男生。在 A2 水平上，达到该水平的女生更多，为 25%，男生则为 16%。在 B1 水平上，女生人数高于男生，女生为 4%，男生为 3%。然而，在为期一年、实施写作干预的研究结束之后，出现了显著进步：共有 90 名学生提升至 B2 水平，8 名学生提升至 C1 水平。B2 归类为独立使用者，C1 归类为熟练使用者。从性别分布数据可以看出，与男生相比，女生在学年末的进步更为明显。处于 B2 水平的女生为 41%，同一水平的男生为 34%。不过，在 C1 水平上男女人数相等，均为 3%。这意味着，这些学生在其后续的大学学习中很可能取得成功。

学年之初实施的第一次学术英语写作测试与学年末实施的第二次学术英语写作测试所得的发现，均运用扎

根理论方法论加以归类。这些板块包括学术英语写作规范、语法规则运用的不一致、词汇与句法结构的使用。在这些主题之下,又形成了其他范畴,以便对实证数据作出更好的诠释。在学术英语写作规范方面,讨论涉及写作风格、内容缺失与写作冗余。在考察语法规则运用不一致这一板块时,形成了三个主题,包括时态的使用、动词的非常规使用以及词汇范畴。随后,在词汇与句法结构使用板块,提炼出句子中的同音异义词与语法规则疏漏两个主题。基于上述发现可以看出,中学与大学的学术英语课程中都需要实施更多与语法规则相关的写作任务。

研究的最后一部分考察了来自学术英语语言测试的访谈数据、对两次测试的反馈以及写作干预情况。数据来自 30 次访谈的编码,用以呈现参与者的声音。学生对学术英语语言测试的反馈中浮现出四个主题:从学术英语测试中获得了信心、测试是有益的、写作能力有所提升、学术英语测试具有挑战性。这印证了测试的运用对学生是有利的,因为他们发现测试最终提升了自身学术英语写作能力的水准。此外,学生对写作干预任务的反馈中浮现出三个关键主题:写作任务带来了进步、学到了新技能、任务增强了意识。高等教育机构可以把这些写作任务纳入其一年级学术英语课程,因为学生认为这些任务颇有助益,并能够借此建构自身的学术英语写作能力。这些能力对其在三年制大学课程中的进阶至关重要。从学生的视角理解学术测试与写作干预具有关键意义。这样做能够为我们提供参考,据以在高等教育中推行新的政策,以促进教育平等并提供知识可及性。

## 八、结论

本研究因其所运用的方法论取向而在语言测试领域具有重要意义。在纵向研究中运用扎根理论取向,这在语言测试中颇为罕见,将拓展当前语言测试研究的开展方式。此外,在第一学年之后建立语言评估机制所具有的教学价值,也将有助于各大学据以设计一年级课程。值得注意的是,本研究尝试建立语言能力测试的制度化与斐济高等教育体系之间的关联。目前,语言能力测试属于自愿性质,且主要用于国际移民目的。因此,本文讨论了把 CEFR 应用于本科生写作能力研究的情况,以衡量其在非欧洲多语情境中的适切性与实用性。本文首先讨论了 CEFR 的适切性,同时对 CEFR、其量表以及学生在量表上的评定结果作了简要评述。文献调研表明,扎根理论方法论(GTM)在语言测试领域鲜有应用。然而,本文通过阐明必要步骤、数据收集工具、数据分析以及呈现结构,说明了扎根理论方法论如何能够运用于语言测试研究。

此外,本文有助于引导扎根理论方法论的初学者将其应用于语言学领域,尤其是语言测试研究。本研究的方法论贡献及其独特的数据集,将推动学界与社会政策层面围绕这一议题的讨论。本研究就扎根理论研究方法如何应用于纵向语言测试研究情境这一问题,对知识体系作出了原创性贡献。本文对硕士与博士研究生等新手研究者以及政策制定者将该方法论应用于自身研究具有借鉴意义。

## 参考文献

- [1] Alderson J C, Figueras N, Kuijper H, et al. Analysing tests of reading and listening in relation to the Common European Framework of Reference: The experience of the Dutch CEFR construct project. *Language Assessment Quarterly: An International Journal*, 2006, 3(1): 3-30.
- [2] Callies M, Diez-Bedmar M B, Zaytseva E. Using learner corpora for testing and assessing L2 proficiency. *Measuring L2 proficiency: Perspectives from SLA*, 2014, 78(71): 79-80.
- [3] Chandra S, Mohammadnezhad M. Investigating factors influencing patient trust in a developing Pacific Island Country, Fiji, 2018. *Heliyon*, 2020, 6(12): e05680.

- 
- [4] Diez-Bedmar M B. The Use of the Common European Framework of Reference for Languages to Evaluate Compositions in the English Exam Section of the University Admission Examination. *Revista de Educacion*, 2012, 357: 55-79.
- [5] 吴莎. 《中国英语能力等级量表(2024版)》新增量表的效度验证. *现代外语*, 2025, 48(4): 525-535.
- [6] 齐林意, 邸爱英, 刘欣刚, 等. 基于《量表》的英语口语测试开发及效度验证研究. *中国考试*, 2024(5).
- [7] Fennelly M G. The influence of CEFR on English language education in Japan. *Bull Shikoku University Online Repository*, 2016, 46: 109-122.
- [8] Goundar P R, Bogitini L J. First Year Fijian Undergraduate University Students Academic Essay Writing Error Analysis. *International Journal of Applied Linguistics and English Literature*, 2019, 8(1): 172-177.
- [9] 彭川, 刘建达. 《中国英语能力等级量表》与《欧洲语言共同参考框架》阅读技能级别对接研究. *中国考试*, 2023(8): 55-64.