

导 读

1. 引言

在应用语言学研究领域，语言测评研究往往会给人留下以定量分析为主的印象。这主要与语言测评研究的侧重点相关。长期以来，考试的效度研究一直是语言测评研究的重点，传统的效度研究模式强调对考试数据进行定量分析。近些年，混合研究方法在语言测评研究领域得到越来越广泛的应用 (Turner, 2014)。所谓的混合研究方法指的是在一项研究中同时采用定量和定性两种研究方法，并将定量和定性研究结果进行整合，从而能够科学、全面地理解相关研究问题。

Jang *et al.* (2014) 指出混合研究方法在语言测评研究中之所以得到越来越多的关注，有两个重要原因。其一，它与语言测评的对象也就是语言能力相关。我们知道，要科学、有效地测量语言能力，就必须先要对其进行定义，而对语言能力和语言使用进行定义困难重重。交际语言能力涉及多个组成部分且各部分之间的关系复杂，而语言使用受到众多因素的影响。因此，单靠定量或者定性研究难以揭示语言能力和语言使用的复杂性。混合研究方法综合了定量和定性研究的特点，比单纯使用其中一种研究方法更能帮助研究者有效地理解语言能力和语言使用的本质。其二，近些年测试效度理论研究取得了重要进展。传统的效度研究忽视了考试的社会学属性 (McNamara, 2001)。Messick (1989) 指出，考试的效度研究不仅要关注考试是否能够客观

地体现考生的语言能力，同时也需关注考试的价值判断和后果。近些年兴起的效度论据观认为，效度研究的过程就是收集效度证据并将其组合形成效度论据的过程。在此背景下，混合研究方法在语言测评领域的兴起具有其必然性。Jang *et al.* (2014) 对近些年在国际期刊上发表的相关论文进行系统分析后发现，采用混合研究方法的研究在语言测评领域呈现较快的增长态势。

尽管混合研究方法在语言测评研究领域得到越来越广泛的应用，但是在本书出版之前，没有专著详细探讨过混合研究方法在语言测评中的应用问题。在教育研究领域，已有多本著作出版探讨混合研究方法的理论与实践（如 Creswell, 2013），而且还有专门聚焦混合研究方法的学术期刊 *Journal of Mixed Methods Research*；在语言教学研究领域也已有探讨混合研究方法的专著出版（Brown, 2014）。本书的出版填补了系统介绍混合研究方法在语言测评研究中的应用这一空白。本书的三位编者中，两位从事语言测评研究，另一位从事混合研究方法理论与实践研究。这样的组合也体现了本书的宗旨，即探讨混合研究方法在语言测评开发和效度研究中的应用。

2. 各章主要内容

本书共分为四个部分，其中第一部分聚焦语言测评与混合研究方法结合时要考虑的基本因素，第二部分探讨开展混合研究方法研究的重要组成部分、主要步骤和考虑因素，第三部分展示语言测评领域采用混合研究方法的个案研究，第四部分提出混合研究方法在实际应用中的建议。下面对各章内容进行简要介绍。

第一章

本章主要介绍混合研究方法在语言测评研究领域应用的背景。首

先，作者讨论了语言测评学科的性质和关注的主要问题，然后讨论该学科近些年的发展趋势。作者指出，语言测评正在从测试文化（testing culture）向测评文化（assessment culture）转变。测试文化注重对语言考试心理测量学属性的探讨，例如考试的信度、效度等；测评文化注重测试与教学之间的关系，强调语言测评的政治、社会、教育和伦理学等方面的属性。这一发展趋势对语言测评领域具有重要意义，也是混合研究方法在该领域应用的重要学科背景。同时，交际语言教学与测试尤其是任务型教学与测试的兴起也是混合研究方法应用的重要背景。

作者在本章列举了混合研究方法的多项显著优势，并以此论述在新的形势下语言测评研究领域采用混合研究方法而非单纯的定量或定性研究的必要性。混合研究方法的显著优势包括：采用多种不同的方法对同一问题进行深入探讨、提升研究设计的科学性和合理性等。例如，定性研究结果可以为后期开展定量研究提供参考；定量研究结果可以为后期的定性研究奠定基础等。作者指出，研究人员可以通过采用不同的混合研究方法设计对语言测评中的多种研究问题展开深入探析，并可以借此拓宽或深化对相关问题的研究；同时，从多个视角对相关问题和观点进行探讨也有助于促进学术交流。

第二章

本章在前一章的基础上继续讨论混合研究方法在语言考试开发、实施和效度研究中的应用。作者首先借助于 Weir (2005) 提出的社会认知效度框架 (socio-cognitive validity framework) 探讨语言能力构念和考试的效度研究问题。该框架从认知和环境两个角度构建语言测试效度研究的理论模型，已在多项国际语言考试（如雅思考试）的效度研究中运用，在国际上具有重要影响。作者通过对该框架的介绍并以本书中的个案研究为例介绍混合研究方法在考试效度研究中所发挥的重

要作用。本章作者 Nick Saville 具有多年语言测试开发和研究经验，善于从考试开发的角度探析混合研究方法的作用。作者指出，语言考试的开发是一项系统工程，涉及多个不同的步骤。为了确保语言考试的质量，语言考试开发机构应在考试开发的过程中采用质量管理体系 (quality management system)，混合研究方法可以为考试开发各个阶段的设计和决策提供重要证据，从而确保决策过程的科学性和合理性。前两章为本书的第二、三部分提供了重要背景介绍，有助于读者从宏观上把握混合研究方法在语言测评研究中的应用。

第三章

本章的作者之一 John Creswell 专门从事混合研究方法理论与实践研究多年，著作等身，在国际上具有重要影响。本章聚焦于混合研究方法的一些基本问题，包括混合研究方法的定义、混合研究方法作为一种新的研究范式的发展历程、混合研究方法的价值观和哲学观、开展混合研究方法研究所需要的技能、提升混合研究方法研究效度的策略以及在语言测评领域开展混合研究方法研究的建议。作者指出，混合研究方法研究并非单纯地收集定量和定性数据并将这两种数据合并；更为重要的是，混合研究方法需有效综合这两种数据从而产生新的研究结果，而这些结果单靠定量和定性数据均无法实现。定性研究人员通常持建构主义的哲学观点，定量研究人员通常持实证主义的哲学观点。这两种观点在其所代表的本体论、认识论和方法论上均具有显著差别。那么，这两种观点如何能够在混合研究方法中共存？作者指出，混合研究方法研究人员有可能持有某种哲学观点，也有可能持有多种不同的哲学观点，或者在某项研究中根据其所采用的混合研究方法设计采用某种哲学观点。研究人员在混合研究方法研究中应清晰表述其持有的哲学观点。除此之外，研究人员应采用多种措施确保混合研究方法研究的效度。例如，研究人员应在研究设计中清楚展示混合研究

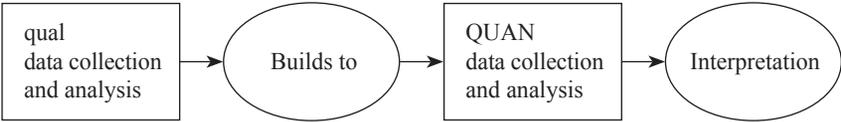
方法的研究步骤，同时也有必要参照国际上通行的研究评价标准和混合研究方法的良好行为规范开展相关研究。

第四章

基于前一章对混合研究方法基本问题的介绍，本章着重探讨混合研究方法的设计问题。首先，作者介绍了混合研究方法的组成部分。作者指出，研究人员在采用混合研究方法的过程中既可能采用归纳法，也可能采用推论法对其研究结果进行解释。如果要从定量和定性数据中获取灵感并进行有效推断，研究人员还需具备第三种思考和推理方法——溯因推理（abduction）。该推理方法介于归纳法和推论法之间，需在这两种方法之间反复，其目的是能够从研究数据中推导出最佳解释。这一方法是混合研究方法研究中定量和定性数据产生协同作用的基础。除了以上三种方法，作者同时介绍了混合研究方法使用中的一个重要概念——元推理（meta-inference）。作者指出，定性和定量方法的有效结合可以产生元推理，这种推理单靠定性或定量方法无法产生。所谓的元推理指的是将通过定性和定量方法得出的推理进行有效组合所形成的整体结论或解释等。作者在本章还介绍了混合研究方法在设计、方法和解释层面的组成部分。在设计层面，混合研究方法可分为五种类型，包括聚合平行设计（convergent parallel design）、解释性混合研究设计（explanatory mixed methods design）、探索性混合研究设计（exploratory mixed methods design）、嵌入式混合研究设计（embedded mixed methods design）和多阶段混合研究设计（multi-stage mixed methods design）。除此以外，作者也介绍了方法和解释层面经常采用的策略。值得一提的是，作者在本章还采用混合研究方法的五种设计类型对近些年语言测评领域中的研究案例进行回顾。最后，作者从设计质量、解释的严谨性和推断的可推广性等角度探讨了混合研究方法在实际应用中的效度问题。

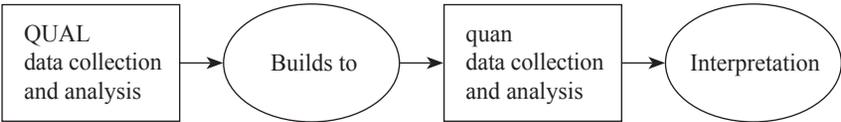
第五章

由于混合研究方法的设计一般比较复杂，因此研究人员经常需要采用图示的方法呈现其研究设计，这正是本章的重心所在。本章首先介绍了混合研究方法步骤流程图的基本规范和制作步骤，然后以不同的研究设计和目的为例展示混合研究方法的步骤流程图。在此，我们以比较简单的探索性混合研究方法设计为例展示混合研究方法的步骤流程图。作者指出，即使采用相同的混合研究方法设计，但是如果研究目的不同，那么步骤流程图也不相同。例如，如果研究目的是为了探讨定性研究的相关结果是否适用于更大的样本，那么研究人员可能会先采用定性研究，然后采用定量研究方法。其简单的步骤流程图如下：



在该图中，小写的 qual 代表定性研究在整个研究设计中处于从属地位，大写的 QUAN 代表定量研究处于主导地位。图中的箭头表示研究步骤，即该研究始于定性研究，根据定性研究结果设计定量研究部分，最后研究人员综合分析确定定性部分研究结果在较大的样本上是否具备可推广性。

同样采用探索性混合研究方法设计，但是如果研究目的是为了对定性研究结果进行归类，那么步骤流程图会有所差异：



在该图中，大写的 QUAL 代表定性研究在整个研究设计中占据主导地位，小写的 quan 代表定量研究为从属地位。和前一个例证相似，定量研究的设计基于定性研究结果，最后研究人员通过综合分析，对定性研究结果进行归类。

在本章中，作者展示了多种不同的混合研究方法步骤流程图，包括一些比较复杂的混合研究方法设计，如嵌入式聚合混合研究方法个案研究设计。本章的另一个重点是教会读者如何采用制图软件来制作混合研究方法步骤流程图。作者介绍了几种常用的制图软件，然后以 PowerPoint 为例详细展示了步骤流程图的制作方法，包括如何在图中展示定量和定性数据的收集和分析方法、如何为流程图添加标题和各个步骤的时间计划等。同时，作者也分享了制作流程图的一些实用建议，如流程图的呈现应尽量简洁、清晰；即使是复杂的混合研究方法设计，流程图也应尽量在一张纸上呈现等。本章对于计划采用混合研究方法的研究人员非常实用，尤其是步骤流程图的制作方法部分。读者可以根据拟采用的研究设计，按照本章中介绍的方法制作步骤流程图并将其呈现在研究论文或报告中，方便读者快速了解研究设计，准确解释研究结果。

第六章

对于计划采用混合研究方法的语言测评研究人员而言，第六章的内容同样非常实用。在本章中，作者论述了撰写和发表混合研究方法研究论文的十大关键要素，并对每个要素进行了详细论述，包括如何为文章拟定标题、如何撰写混合研究方法研究论文的开头部分、如何清晰地撰写研究目的、如何拟定研究问题、如何介绍研究方法的哲学基础、如何介绍研究中所采用的研究方法和步骤、如何有效呈现研究结果、如何对研究结果进行讨论，甚至包括参考书目和附录的准备等。这十大要素包括很多非常实用的具体指导和建议，相信读者和笔者一

样，在读完之后感觉受益匪浅。例如，作者在论述第四个要素，即如何拟定研究问题时指出，定量部分的研究问题宜以研究假设或问题的形式呈现，而定性部分的研究问题宜采用开放性问题的形式，在语言表达上宜采用中立、不带方向性的语言。除了以上这些要素以外，作者还为读者提供了文章发表方面的中肯建议，如期刊的选择、在准备稿件时的注意事项等。除了这些实用性的建议之外，本章还有另外一个特色，那就是本章的附录。读者可以根据在本章中所学习到的内容尝试完成附录中的练习，这有助于吸收本章的相关知识，巩固学习成果。

第七章

本章的重点是混合研究方法在语言测评效度研究中的应用。作者首先探讨测试效度的定义，指出效度包含多个不同的方面，效度研究就是系统收集效度证据并将其进行组合的过程；同时，效度研究也是持续不断的过程。如前文指出，混合研究方法最大的优势在于比单纯的定量或定性研究能够更深入、更全面地理解研究问题。为了论证混合研究方法在效度研究中的重要作用和所面临的挑战，作者收集并分析了语言测评研究领域中采用混合研究方法的 10 篇论文。作者指出，混合研究方法在效度研究中的应用可能面临重要挑战，包括：如何综合定性和定量部分的研究结果、如何确保定性和定量数据收集和分析的严谨性、如何有效呈现定性和定量部分的研究结果、如何讨论采用混合研究方法为研究增添新的价值等。作者逐一论述了这些挑战，并强调混合研究方法对于开展语言测评效度研究可能发挥的重要作用。以此，作者得出了本章的结论，即混合研究方法对提升语言测评的效度和构建其效度论据具有重要作用，因此语言测评研究人员应更多地采用混合研究方法提升考试的效度。

第八章

本章到第十一章介绍采用混合研究方法的个案研究，以实例展示混

合研究方法在语言测评开发和效度研究中的重要作用。本章以国际法律英语考试 (The International Legal English Certificate) 为例展示混合研究方法在考试开发中的应用。作者首先介绍国际法律英语考试开发的背景, 然后介绍该考试开发项目中所采用的独特的研究设计——社区参与研究 (Community-based Participatory Research)。与传统的研究设计不同, 社区参与研究强调语言测评专家与非学术研究人员如法律专家之间建立合作关系, 提升考试内容和设计的效度。然后作者介绍了在国际法律英语考试的开发过程中如何采用社区参与研究和探索性混合研究方法设计, 确保考试开发过程的科学性和严谨性。

该考试开发的第一阶段主要采用定性研究设计进行需求分析 (needs analysis); 基于第一阶段的研究结果, 第二和第三阶段主要采用定量研究方法。例如, 在第二阶段, 研究人员采用问卷调查收集、研究考试的利益相关群体 (如考生和法律英语教师) 对考试效度和考试是否适应市场需求等方面的看法。基于研究结果, 考试开发人员对考试的内容和形式进行修订。在第三阶段, 研究人员采用经典测试理论和项目反应理论对考试数据进行定量分析, 并根据分析结果对考试进行再次修订。基于混合研究方法的研究为考试开发的不同阶段提供了多种类型的重要证据, 从而确保考试开发的科学性。本章内容对于语言测评开发人员尤其是专门用途英语 (English for Specific Purposes) 考试的开发人员和机构具有重要的参考价值。

第九章

对于口语测评而言, 评分量表的开发和使用直接影响考试的信度和效度。本章聚焦于混合研究方法在高风险口语考试评分量表开发中的重要作用。作者指出, 对于大规模口语考试而言, 评分量表的开发考虑因素众多, 是非常复杂的工作, 而混合研究方法可以从研究方法层面为量表的开发提供重要支持。本章展示了较为复杂的混合研究方法

设计在口语评分量表开发中的应用。在该项目中，作者首先采用定性研究，然后采用定量和定性平行的同步研究设计，最后采用定量研究。在研究的第一阶段，作者对学习者的语篇进行话语分析；同时，作者还邀请了评分员对已有的评分量表进行评论。基于本阶段的研究结果，作者开发出口语评分量表的初稿。在研究的第二阶段，作者采用有声思维的方法探讨评分量表的使用问题；同时，作者也采用多层面 Rasch 模型对评分员的评分结果进行定量分析。基于本阶段的研究结果，作者对评分量表进行修订。在研究的第三阶段，作者主要采用多层面 Rasch 模型探讨评分量表在实际评分中的应用问题。基于本阶段的研究结果，作者确定了评分量表的终稿。尽管本章探讨的对象为高风险口语考试，但是混合研究方法同样适用于其他表现型语言考试（如写作、翻译）评分量表的开发与效度研究。本章另外一个重要启示是从事混合研究方法研究需要多项研究技能。以本章研究为例，研究人员需具备话语分析、有声思维等定性研究技能，也需要具备多层面 Rasch 模型等定量研究技能，这对于研究人员无疑是重大挑战。本书的最后一章将探讨如何应对这些挑战。

第十章

本章聚焦于混合研究方法在阅读测试开发与修订中的作用。本章所涉及的语言考试是剑桥大学考评部开发的一项针对高水平英语学习者的阅读测试，主要用于大学入学选拔。作者在本章中主要介绍了采用混合研究方法对该考试的内容和形式进行修订的过程。由于考试的修订涉及多个步骤，因此与前一章相似，本章采用了比较复杂的混合研究方法设计以解决在不同步骤中出现的问题。本章中的研究共分为三个阶段，分别为阅读任务设计、试测和效度论证。在任务设计阶段，作者主要采用定性研究。该阶段研究的主要目的是了解目标语言使用域对于学生的阅读能力需求，以此确定阅读测试的内容和形式。基于

该阶段的研究结果，作者确定了考试的任务类型。在试测阶段，作者采用了定性和定量平行的同步研究设计。定性研究主要包括专家判断，即邀请阅读专家对考试内容和形式发表评论；定量研究主要包括采用经典测试理论和 Rasch 模型对考试数据进行分析，探讨试题的质量。该阶段的研究结果对考试任务的修订提供了重要支撑。在效度论证阶段，作者主要采用定量研究，包括采用 Coh-Metrix 对阅读文本进行语篇分析、采用经典测试理论和 Rasch 模型对考试数据进行分析等。该阶段的研究结果为阅读测试提供了重要的效度论据。正如作者在本章的结尾部分所言，本章并非是纯粹的学术研究，而是展示混合研究方法在现实考试项目中的应用。与前两章相似，本章对于语言考试开发人员和机构具有重要的借鉴意义。

第十一章

自 Messick (1989) 提出其效度理论以来，考试的影响和后果研究成为语言测评效度研究的重要组成部分。本章探讨的重点为如何将混合研究方法应用于考试的影响研究。作者指出，考试的影响研究涉及到多个不同的考试利益相关群体，如学生、教师、家长等，且各个群体关注的重点也各不相同。由于考试影响研究的涉及面和关注点众多，因此单靠定性和定量研究很难完成。混合研究方法综合定性和定量研究的优点，在考试影响研究中具有独特优势。作者以一项国际语言考试（剑桥少儿英语考试）为例，采用定量和定性平行的同步研究设计，探讨该考试对越南英语学习者所产生的影响。定量研究部分包括大规模问卷调查和学生考试数据分析，定性研究部分包括半结构访谈和核心组讨论。该研究的参与者包括学生、教师、家长、政策制定者等。研究结果包括多个方面，如学生英语学习取得的进步、英语课程的有效性、学生的学习动机、教师和家长对于考试的态度、教师的课堂教学等。作者对不同的群体所产生的研究结果进行比较，从中发现一致

与不一致的地方并对其解释。而且，作者围绕研究涉及的不同方面展示了定量和定性两种研究结果。本章对于计划开展语言测评影响研究的科研人员具有重要的参考价值。

第十二章

本章为本书的最后一章，主要是对本书的内容进行总结，并且给计划在研究中使用混合研究方法的研究人员提供使用建议。作者共提出了十条建议，并且对每条建议进行了具体阐述。这十条建议可分为三个部分，分别为：1) 采用混合研究方法设计时所需要考虑的基本要素；2) 混合研究方法研究问题的提出和研究设计；3) 混合研究方法的数据收集和分析以及研究结果的呈现。在阐述每条建议的时候，作者先提出问题，然后提供解决方案，最后通过本书中的研究案例对问题和方案进一步阐述。前文已经指出，混合研究方法，尤其是复杂的混合研究方法设计，对研究人员的研究技能有很高的要求。如果研究人员不具备这些技能，就很难确保混合研究方法研究的严谨性。作者指出，如果计划采用混合研究方法开展研究，但缺乏相关的技能，那么研究人员一方面可以考虑拓展自己的研究技能，另一方面也可以考虑和其他研究人员建立合作关系，共同开展研究。这正是作者提出的第一条建议。在每条建议讨论结束以后，作者还提供了一些实用的操作步骤，帮助读者应对研究中的挑战。在本章的最后一部分，作者号召读者积极行动起来，在研究中尝试采用混合研究方法以解决语言测评领域的多种研究问题。

3. 结语

目前，混合研究方法在我国应用语言学研究领域的应用很少。孟春国、陈莉萍（2015）将我国外语研究领域发表的论文与国际应用语言

学领域同期发表的论文进行比较后发现，定量研究在国内期刊论文中占据绝对主导地位（占 87.5%），定性研究数量稀少（占 3.57%）；国内期刊刊载的混合研究论文数量偏少，尤其是质量上乘的研究范例数量更少。江进林（2018）对我国外语测试界在 2006–2017 年期间所开展的研究进行详细分析后发现，我国的外语测试研究涉及效度、反拨效应、评分等多个主题。从研究方法上来看，大部分研究采用定量或定性的研究方法，运用最多的是问卷调查与访谈，但是将定量和定性研究有效综合的混合研究设计依然少见。因此，本书对于我国的语言测评研究人员是不可多得的研究方法参考书目。书中既有对混合研究方法和语言测评的理论探讨，也有展示混合研究方法在语言测评开发和研究中实施的案例。相信读者和笔者一样，在读完本书后定会加深对混合研究方法的理解并积极尝试将其应用到自己的研究中。

范劲松

2019 年 3 月 20 日

参考文献

- Brown, J. (2014) *Mixed Methods Research for TESOL*. Edinburgh: Edinburgh University Press.
- Creswell, J. W. (2013) *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. Thousand Oaks, CA: Sage.
- Jang, E. E., Wagner, M. and Park, G. (2014) Mixed methods research in language testing and assessment, *Annual Review of Applied Linguistics* 34, 123–153.
- McNamara, T. (2001) Language assessment as social practice: Challenges for research, *Language Testing* 18(4), 333–349.
- Messick, S. (1989) Validity, in Linn, R. L. (Ed) *Educational Measurement* (3rd edition). New York: Macmillan, 13–103.

Turner, C. E. (2014) Mixed methods research, in Kunnan, A. J. (Ed) *The Companion to Language Assessment* (1st edition). Hoboken: Wiley & Sons, 1, 403–1, 417.

Weir, C. J. (2005) *Language Testing and Validation: An Evidence-based Approach*. Basingstoke: Palgrave Macmillan.

江进林. (2018). 我国外语测试实证研究: 回顾与展望——基于外语类主要期刊的统计分析(2006—2017). 《外语界》, 185(2), 40–48.

孟春国, 陈莉萍. (2015). 走向多元融合的研究范式——中外应用语言学与外语教学期刊的载文分析. 《外语界》, 166(1), 2–11.