

实证研究中的文献述评的立体功能框架： 以语文综合评估研究为例

The Framework for Literature Reviews with Multiple Functions in Empirical Papers: Taking Chinese Language Integrated Assessment Research as an Example

祝新华 陆琦

Zhu Xinhua Lu Qi

【摘要】适当的文献述评是确保实证性论文具有原创性的基础。然而，过往的一些文献述评明显存在有待改进之处，如“多述而少评”“过度评价”和“述评失焦”等。基于两篇语文综合评估的实证研究实践，本文建构了“三阶段*两层次”的文献述评功能框架。三个阶段指“整合”“评价”与“提问”，两个层次包括“奠基”和“创新”。“三阶段”与“两层次”交互作用，形成文献述评的立体功能。在此基础上，本文初步提出达成这些功能的策略，可为研究者写好实证研究的文献述评提供参考、借鉴。

【关键词】实证论文，文献述评，立体功能，达成策略

Abstract: A well-organized literature review (LR) is the basis for the originality of empirical papers. However, there are some issues in literature reviews in existing studies. To address these issues, this paper, taking two empirical studies in Chinese language assessment for examples, constructs a functional framework for LR, including “three phases and two levels”. The three phases consist of “integration, evaluation and questioning” in

祝新华，香港理工大学中文及双语学系。

陆琦，浙江大学教育学院。本文通讯作者。

本文是香港General Research Fund (GRF) 研究项目 (No 15640416) 的部分相关成果。

LR writing, while the two levels refer to the “fundamental and innovational” levels. The three phases and two levels integrate to form a dynamic framework of LR with multiple functions. The strategies to realize those functions are expounded. This paper aims to provide researchers with insightful ideas for LR writing.

Keywords: empirical papers, literature review, multiple functions, improvement strategies

一 实证研究中的文献述评

实证研究基于事实和证据,通过质性或量化的研究方法,引导进行关于问题的逻辑推理,从而探寻关于研究对象“可能是什么”的相关研究(Hansen & Gerst, 1967; National Research Council, 2002)。实证研究经由自然科学研究的发展,目前已成为人文、社会科学研究的重要范式之一。因此,人文、社会科学的国际学术期刊也多发表实证研究的论文。这种论文一般由概述、文献述评(literature review)、方法、结果、讨论、结论等部分组成。它旨在通过文献述评找到已有研究的不足,同时联系相关现状提出有意义的研究问题,采纳已获认可或经验证明的研究工具以取得第一手数据,使用合适的分析方法以总结有价值的研究结论。问题的前沿性与结果的原创性是此类研究的重要特征,而适当的文献述评则为确保这种特征奠定基础。

文献述评是对文献进行搜集、整理、分析、评价等的研究活动。它以述评已有研究取得的成果及不足来体现当前研究的合理性,并展现新研究的理论与实践价值。为了揭示文献述评的特点、引导人们进行适当的文献述评,不少学者从步骤(steps)、阶段(phases)或任务(tasks)等角度探讨文献述评的过程。如Machi和McEvoy(2016)的六个步骤:选择主题、发展论点、搜索文献、研究文献、批判文献、撰写述评。Cooper(1998)的五个阶段:问题形成、数据收集或文献搜寻、数据分析、分析与解释、展现研究结果。Fink(2014)的七项任务:选择研究问题、选择文献及数据库来源、确定检索关键词、制定筛选标准、应用标准评估、进行述评、综合述评结果。上述观点体现了文献述评确立主题或生成研究问题的过程。而在文献述评的写作上,“倒三角形法”是较受人们关注的写作方法,它要求先叙述与研究主题关系较远的文献,然后叙述较近的文献,最后带出研究问题(崔建军,2014)。不过,这种方法未能较清楚地引导人们“如何述”与“如何评”。

在中文学术环境中，文献述评有“文献综述”“文献回顾”“文献探讨”“文献评价”等不同的称谓。尽管这与研究者的语言习惯有关，但也正好体现了文献述评存在对“重述”还是“重评”的理解差异。与之相对应，研究者在文献述评时容易陷入“述评失衡”的两种误区：一是“多述而少评”，对已有文献过多陈述，而忽略对文献内容作必要的优劣价值的评判。这样的述评经常止步于文献的列举，缺乏对已有研究科学性的论证。二是“过度评价”，对文献成果提出随意的主观论断。这种论断由于对文献内容及其关系缺乏全面整理，评价没有稳实基础，观点往往有失偏颇，无法构成当前研究所需的学术基础。从文献述评的功能看，还有一种“述评失焦”的误区，即文献述评的过程没有为提出研究问题而服务。它具体表现为：不对文献进行应有的评价以提出研究问题，或提出的问题与述评之间并不存在实质的关联。上述误区均会限制文献述评的质量。

二 文献述评的立体功能框架

本文结合两篇实证研究论文作为实例来建构“三阶段*两层次”的文献述评框架。第一篇论文是Zhu等（2019a）的有关口语沟通策略在综合小组讨论中的应用的研究，题目为*Strategy use in oral communication with competent synthesis and complex interaction*；第二篇论文为Zhu等（2016）的有关语文综合写作与独立聆听能力之间的关系研究，题目为*Exploring the relationships between independent listening and listening-reading-writing tasks in Chinese language testing: Toward a better understanding of the construct underlying integrated writing tasks*。“三阶段*两层次”的文献述评框架，包括文献述评的立体功能及其达成策略。其中，三个阶段包含对过往研究的“整合”“评价”与“提问”的三个思维过程；两个层次则要求文献述评不仅要为新研究“奠基”，还要体现“创新”（见表1）。本文通过建构本框架，力图改进文献述评的实践，从而为进一步提高研究的学术水平作出贡献。

表1：文献述评的立体功能

	整合	评价	提问
奠基	客观总结全球已有研究成果，特别关注其中共同的、不同的，甚至相反的方法、观点、结论与建议。这有助于体现新研究的建构基础。	在审视全球研究成果的基础上，揭示已有研究的进展，特别是其中的不足、空白点。这有助于体现新研究的理论价值。	概括现实场境，即相关的改革或实践、重点留意值得全球关注的现实问题。这有助于体现新研究的实践价值。
创新	导出与新研究直接相关的概念、理论，建构新研究的变量、框架或模型。也可作必要的拓展。	点出研究者的兴趣，揭示新研究的学术需要、动机。	结合已有研究中存在的问题，如不足、空白点与现实问题，提出新研究的问题。如有足够的文献支持时，可提出精细化的研究假设。

2.1 整合

2.1.1 功能

第一个阶段是整合已有研究，它包含两层功能。外层的“奠基”是客观总结已有研究成果，要求研究者找出与研究主题相关的文献，通过概括已有研究在方法、观点、结论、建议方面达成的共识及存在的差异。内层的“创新”则基于对已有研究的整理，导出与新研究直接相关的概念、理论，建构新研究的变量、框架或模型。在具体操作中，概念既可从已有研究中直接引入，必要时也可根据新研究的特点对概念的解读进行有理据的调整或拓展。变量是研究对象的组成因素，指向研究需要考察的维度，常呈现为分析指标。为了取得数据，变量通常要具备可测量的特点。研究框架是对研究对象所包括的因素（如变量）及因素间的相互关系的阐述，它是新研究的建构思路。

以Zhu等（2019a）口语沟通策略的研究为例，研究者在搜寻到与该主题相关的文献材料后，便需要对已有成果进行系统整理。例如，沟通策略可分为两类：心理语言的（psycholinguistic）与交互的（interactional）策略。前者主要关注沟通中个人语言问题的解决，而后者则聚焦于通过意义协商以取得共识（Faerch & Kasper, 1984；Nakatani, 2006）。对于交互策略，根据第一语言的特点，Zhu等（2019a）对后者加入了如下因素以拓展沟通策略的内涵：在小组讨论中通过提供解释或案例，以及评论他人的观点，从而实现对观点的精细加工与整合的目的。

如表2所示，研究者进一步比较了“沟通策略的分类”。表格的横向陈述

了常见的几种沟通策略模型：语言学习策略分类（language learning strategy classification）、语言策略分类（categories of language strategies）、话语生成模型（speech production model）、说话策略类型（types of strategies crucial to speaking）；纵向则根据策略类型梳理了模型所包含的各种策略，比如元认知策略（metacognitive strategies）、认知策略（cognitive strategies）、情感策略（affective strategies），以及其他策略（other strategies）。若某一模型的策略类型显示为空格，表明该模型没有涵盖该类型的策略。由此可见，有关沟通策略的研究主要聚焦于元认知、认知及情感策略，这三种策略在口语沟通中积极作用得到了研究者较为一致的认可。另外，语言策略（language strategies）、补偿策略（compensation strategies）、语言学习策略（language learning strategies）、语言应用策略（language use strategies）、交流策略（communication strategies）等虽亦在其中，但这些其他策略并不常出现于沟通策略的模型之中，可见它们仅得到了少数研究者的支持。

表2：沟通策略的分类

沟通策略模型 策略类型分类 ¹	语言学习策略 分类 ²	语言策略 分类 ³	话语生成 模型 ⁴	说话策略 类型
元认知策略	元认知策略		元认知策略	
认知策略	认知策略 (包括记忆)		认知策略	认知策略
情感策略	情感策略			情感策略
其他策略	补偿策略			
	社交策略	语言学习策略	语言策略	交流策略
		语言应用策略		

(参考Zhu et al., 2019a, 页1165, 1166)

- 1 详见Oxford, R. L. (1990). *Language learning strategies*. Newbury House Publishers.
- 2 详见Cohen, A. D. (2003). The learner's side of foreign language learning: Where do styles, strategies, and tasks meet? *International Review of Applied Linguistics in Language Teaching*, 41(4), 279-291. 以及 Cohen, A. D. (2010). Focus on the language learner: Styles, strategies and motivation. In N. Schmitt (Ed.), *An introduction to applied linguistics* (2nd ed., pp. 161-178). Hodder Education.
- 3 详见Douglas, D. (1997). *Testing speaking ability in academic contexts: Theoretical considerations*. Educational Testing Service.
- 4 详见Huang, H. T. D. (2016a). Exploring strategy use in L2 speaking assessment. *System*, 63, 13-27. 以及Huang, H. T. D. (2016b). Test-taking strategies in L2 assessment: The test of English for international communication speaking test. *Perceptual and Motor Skills*, 123(1), 64-90.

以此为基础，研究者建构了适用于综合小组讨论的策略框架（Zhu et al., 2019a）。从对已有研究的整理总结，到导出新研究所需的概念、理论，建构变量、框架，这是一个逐层递进的过程。值得注意的是，内层的“创新”功能应力求细致，其作品，如变量、框架应易理解，从而促进研究指标、研究工具的发展。

2.1.2 达成策略

达成上述功能的策略可分为两个方面：搜寻与管理文献、梳理与关联文献。

搜寻文献的第一步是广泛检索适用文献。未清楚自身研究兴趣或未确定主题的研究者，可以尝试先写下一个句子，从句子中找出一些关键概念，接着列出与这些概念相关的术语，而后再以这些术语作为关键词，在“谷歌学术（Google Scholar）”“中国知网（CNKI）”等数据库检索文献。

初步提取适用文献后，研究者可从中找出几篇核心文献。它们通常发表在社会科学引文索引（Social Sciences Citation Index，简称SSCI）、中文社会科学引文索引（Chinese Social Sciences Citation Index，简称CSSCI）等高质量刊物，是具有代表性、原创性的研究（key studies）或文献述评。研究者可根据这些文章的参考文献（reference list）查读相关的文献。这是文献的“追溯”方法。另一种方法是“追踪”，即搜寻那些引用过这些核心文献的新近论文。最后，研究者可参照表3精选文献的参考标准。

表3：精选文献的参考标准

参照对象	参照准则
时间	1. 较早的 / 经典的 / 有重要影响的报告 2. 最新的研究报告
主题	1. 与本研究主题十分密切的研究 2. 与本研究主题有些相关但很有启发的
作者	1. 由本领域专家完成的严谨报告 2. 由不同领域的研究者完成的报告
方法	1. 设计严谨的传统方法 2. 采用非常规 / 新的方法
结论	1. 有典型结论的报告 2. 有不同结论 / 不可思议结论的报告

为了便于文献的储存、查核及日后再用，研究者需要做好文献管理。EndNote、Mendeley等文献管理软件可帮助研究者完成这项工作。EndNote是科学

引文索引 (Science Citation Index, 简称SCI) 的官方软件, 主要功能有在线搜索文献、建立文献库、定制文稿及引文编排。后两项功能是与写作相关的常用功能。因为EndNote在安装完毕后便可自主添加到MS Word文件处理系统中, 研究者在需要引用文献时可点击MS Word工具栏中的“Endnote”, 通过选择“Find Citation(s)”插入文献。Mendeley则是Elsevier的产品, 它除了有完备的文献管理功能以外, 也支持与他人进行在线的学术社交活动。此外, Zotero、Docear、ReadCube、NoteExpress等也可用于文献管理。这些工具既便于研究者在写作时随时取用文献, 也有利于研究者在日后修改文章时自动调整文献出处。

选定文献后, 就要对它们的相关内容进行梳理与关联。为此, 首先要在宏观上考虑如何组织文献。以下几种方法较为常用: (1) 时间顺序: 遵循历史发展的路径, 从早期的文献叙述到新近文献; (2) 方法顺序: 把应用同类研究方法的文献组织到一起, 其中, 把本研究将要采用的方法放在最后归纳; (3) 主题顺序: 用与本研究主题的远近关系为序, 先叙述远相关的, 再叙述近相关的, 最后提出研究问题或假设, 从而形成文献组织的“V”型结构。一般来说, 前两种方法比较适合于仅有单一或较少主题的研究。如有多个主题, 宜用主题顺序组织文献。

对于不同文献中的信息, 要特别关注其中具有差异性或独特性的部分。研究设计部分可重视不同研究中的工具、对象、抽样方法、实验步骤等差异, 研究内容则需重点关注不同的观点、结论, 以及研究建议。研究局限性是作者对研究不足的反思, 我们可参考它提出自己的研究问题, 或藉此辨别影响研究质量的可能因素, 从而使自己的研究设计避免不必要的疏忽或错误。

对于不同文献中的具体信息的表述, 可选择两种方式来完成信息的组织。一是“单篇罗列”, 即引述各文献内的信息, 如“学者X(2019)提出……, 学者Y(2018)建议……”等。这种方式通常适用于十分重要、同时又无其他相关文献可采用的情况。在整篇述评中, 这种内容应只占其中的一小部分。二是“跨篇总结”, 即对文献的共同和不同之处进行概括。具体来说, 多个文献的共同之处可表述为“……(学者X, 2019; 学者Y, 2018)”。例如:

A number of studies have compared students' performance in independent and integrated tasks. Some suggested that independent and integrated writing tasks measure a similar (or the same) construct (e.g., Cumming et al., 2005; Gebril, 2006, 2009, 2010; Lee, 2006; Lewkowicz, 1994; Watanabe, 2001); others argued that the two tasks possess somewhat different constructs (e.g., Guo, Crossley, & McNamara, 2013). (Zhu et al., 2016, 页168)

上文指出，有不少研究比较了学生在独立写作和在综合写作任务中的表现。有的研究指出独立写作任务与综合写作任务评核相似（或相同）的能力结构（如：Cumming et al., 2005；Gebril, 2006, 2009, 2010；Lee, 2006；Lewkowicz, 1994；Watanabe, 2001）；而另一些研究则争论这两项任务的能力结构并不相同（如：Guo et al., 2013）。而对不同之处，则可用“有的研究提出了……（学者X, 2019）；而有的研究则提出了不同的观点……（学者Y, 2018）”一类的表述，例如：

Looking at the product of integrated tasks, some studies have found more complicated functional and textual organization in integrated tasks than in independent tasks (Brown et al., 2005); others have revealed that test takers tended to include more ideas, which are often less developed, in their integrated writings than in independent writing responses (Lewkowicz, 1994).（同上）

上文指出，从两种写作任务的作品来看，有的研究发现，综合写作任务相较于独立写作任务，有着更为复杂的功能与文本组织（Brown et al., 2005）；而另一些研究则指出，综合写作任务的应考者虽然会在写作中包含更多的观点，但这些观点往往没有应有的展开（Lewkowicz, 1994）。

从当前研究现状看，研究者宜少单篇罗列，多跨篇总结。跨篇总结能让读者更快地把握文献之间的内在联系，从而为提出新的问题奠定基础。

2.2 评价

2.2.1 功能

文献述评的第二个阶段是评价已有的研究。它的外层“奠基”功能要求研究者在整合文献的基础上，把握已有研究的进展，并特别关注研究的不足，聚焦可待改进之处。

以Zhu等（2019a）口语沟通策略的研究为例，下述呈现了文献述评应如何在把握现有的研究发展的基础上，进而关注不足：

We identify two issues in previous studies on strategy use during speaking tasks. First, almost all of the studies involved single direction (vs. multiple direction) speaking. That is, they involved individuals answering questions, describing pictures, responding to questions with information provided, proposing solutions and expressing opinions. No study on group discussion among student peers has been conducted.

Therefore, the strategy use elicited from the tasks in previous studies does not reflect the interaction competence between peers. Second, few of the studies involved integrated speaking tasks. A few studies used relatively simple and short passages, talks, lectures or conversations. Furthermore, processing materials in tasks primarily involved combining information or summarising key ideas, rather than synthesising information. Nevertheless, in a real-life context in this globalised era, the requirement for language proficiency is much more complicated. For instance, processing multiple texts with controversial or complementary information, providing one's opinions and arguing them based on synthesised information. Thus, it is important to investigate how undergraduate students demonstrate strategy use in authentic tasks with appropriate source materials provided. (页1167)

从以上段落可见，研究者提出了过往研究存在两个方面的不足：（1）过往研究均关注学生单方向的发言，较少重视学生在小组讨论中与同伴的互动过程；（2）过往研究提供的任务都比较简单，较少涉及真实的综合性任务。所谓综合性任务，是一种体现真正生活情境、涉及综合语言应用能力的复杂任务，如处理具有争议性的或包含互补信息的多个文本，基于信息整合提出具有辩证性的观点等。新研究则对这些不足作出改善，即体现学生在小组讨论中的复杂互动，并考察学生在综合任务情境中的信息整合能力及策略应用情况。

此外，关注已有研究结果的不一致之处也有助于进行多角度的思考。仍以口语沟通策略的研究为例，不少过往研究认为沟通策略与学生的说话水平显著相关，但新近研究则发现沟通策略的数量与学生的说话水平没有显著关系（Barkaoui et al., 2012; Swain et al., 2009）。这种结果差异至少能给予我们两种启发：（1）影响研究结果不一致的因素有哪些？（2）我们能否找到具有显著正面作用的策略以更有效地指导学生提升说话能力？（Zhu et al., 2019b）

在找出已有研究的参考价值与意义后，点出学术研究的新需求或兴趣是评价阶段的内层“创新”功能。如Zhu等（2016）在综合写作的研究中指出，过往的相关研究主要从学生的任务表现来比较独立写作与综合写作的最后作品的异同，并做出了下述评价：

These studies seemed to suggest that while students' performance may be similar across independent and integrated tasks, there are differences in their underlying constructs. However, these findings do not illuminate the relationships between the underlying constructs of the independent and integrated tasks. (页168)

研究者认为,已有研究虽发现了学生在这两种任务上的表现的相似性,但不能解释这两种任务的结构有怎样的关系,这说明该问题还存在继续探讨的空间。

通过进一步的文献梳理,研究者认可了已有研究以英文作为第二语言方面的丰富成果,而有关第一语言特别是中文的国际学术研究很少,并点出新研究的学术需求或兴趣——关注以中文为第一语言的学生的表现,同时期望第一语言的相关研究能为第二语言的教学、学习及评估提供更多有效的信息,促进已具有高语言水平的学生的进一步发展:

Existing research in the field of integrated tasks has also shown that the majority have been conducted in the context of English as a second language (ESL). Few have looked at the testing of English as a first language, but almost none has dealt with Chinese Language testing.....Therefore, we are interested in investigating first language test takers because findings from first language research can inform the teaching, learning, and assessment of the language concerned as a second language, especially for those students with high levels of proficiency. (Zhu et al., 2016, 页169)

简言之,评价作为整合与提问阶段的中间过程,承担“承前启后”的过渡作用。一方面,它是对已有文献价值的进一步解读,通过作出价值判断以增进研究者对已有研究的理解;另一方面,在“问题导向”的文献述评中,评价更为提出新的前沿问题提供思路。

2.2.2 达成策略

文献评价应紧密联系内容,保持中肯适度批判。

第一,体现评价与整合、提问阶段的关联性,即评价要基于对文献的充分搜集与整合。研究新手可参考如下建议:在每个主题叙述完结后,及时跟随评价,呈现“主题+评论”的结构。这种一个主题一个评价的对应式表述,比完成所有的文献列举后再来评价更容易把握。此外,文献中很多内容都可以评价,而选择焦点来进行评价,则有助于逐渐引出研究问题。

第二,客观中肯。文献研究是开放的学术思考,在肯定他人研究价值的同时,也鼓励研究者对文献加以评判。这一过程概述为:保持客观、言辞中肯;批判有据、态度谦逊。在指明他人研究的空白点时,要合理判断,留有余地。如“就我们所知,这是关于XX主题的首次研究……”。另外,任何研究都是在一定的条件下进行的,评价时不能不顾当时条件,提出过于苛刻的要求,而是应思考当时研究中存在的可控或不可控因素。如有研究误差,则思考该误差是否是其研究条件下允许发

生的。这样，研究者既能不偏离评价重点，也有助于思考自己的研究是否面临类似的制约，以及是否可采用可行的应对办法。以综合写作的研究为例，Zhu等（2016）发现了已有研究在学生独立聆听能力测量方面的不足，继而考虑了处理这一“不足”的可行办法：

However, it should be noted that the source material used in the study by Frost et al. was only a single short recording from a radio programme, which may not be sufficient to measure test-takers' listening competence. We are interested here in identifying the cognitive processes involved in listening, how they affect the written product of integrated tasks and to what extent the listening competence of test takers can predict their performance in integrated writing.（页168）

过往研究对学生独立聆听能力的测量不足主要体现在测量方式上比较单一，忽略了学生在聆听任务中的认知参与。因此新研究会完善对学生认知过程能力的分析，探讨它是如何影响学生的综合写作的。

李泉鹰（2011）指出，只有带有批判性或解释性的文献综述才能为“学术增值”。带着研究问题、客观适度地对过往研究成果作出评判将会为已有文献赋予新的学术价值。

2.3 提问

2.3.1 功能

文献述评的第三个阶段是“提问”。它的外层“奠基”功能是：在整合与评价已有研究成果的基础上，结合相关的现实场境提出研究问题。现实场境是研究者熟悉的场境，也是本研究的取样之处，它要与述评中提出的理论研究的不足、空白点有关。因而，这一功能可分为两个操作步骤：一是概括相关的实践或改革现状，二是找到有国际意义的现实问题，如国际上一些学者、教育实践人员关心或正在解决的问题。以Zhu等（2019a）口语沟通策略的研究为例，研究者先概述了香港学生的普通话学习现状。香港回归后虽然实行“两文双语”政策，但是因人们在日常会话中较少使用普通话（Census and Statistics Department, 2016）、香港中小学多以广东话或英语作为教学语言等原因，所以普通话在香港可视为第二语言（Li, 2017）。普通话的学习与应用不足使得普通话的教学效果成疑。与此同时，研究者亦发现，在日常工作或生活中，人们总需要在聆听或阅读一些材料后才做出口头报告或与他人展开讨论。这种综合性的交际过程中所涉及的观点交流、信息整合等能力是全球

现代社会人才必备的综合技能（Zhu et al., 2019a）。可见，研究小组讨论中的沟通策略，既有助于解决香港现实场境中的难题，亦具有国际意义。这构成了这项研究的实践价值。

提问阶段的内层“创新”功能是提出研究问题、假设。在导出这一功能时，除了要从现实场境中的现状问题推进，还要注意连贯前两个阶段的内层功能，即结合新研究的框架及研究者的研究需求，从而提出新研究的问题或假设。例如：

We developed and validated a strategy inventory of integrated group discussion and then examined strategy use and its relationship with task performance among undergraduate students. The following questions were proposed: (1) Which strategies have significant effects on the integrated group discussion performance of undergraduate students? (2) What is the relationship between strategy use and task performance across the concerned strategies and participants' discussion performance levels? (Zhu et al., 2019a, 页1168)

上述研究以本科学生为研究对象，旨在探讨综合小组讨论中应用的沟通策略，因此主要提出了两个研究问题：（1）哪些策略能够显著提升学生在综合小组讨论中的表现？（2）沟通策略的应用与学生在综合小组讨论任务中的表现有怎样的关系？

2.3.2 达成策略

现实场境中的问题构成了研究问题的雏形。在此基础上，研究者可使用策略来帮助发展研究问题。Booth等（2008）曾提出界定研究问题的三步骤，即“正在研究的题目是什么”“想要找出的问题是什么（what）/为什么（why）/如何（how）”以及“该问题的意义是为了帮助读者了解什么”。在上述案例中，现状问题——普通话交流困难，仅能满足其中的第一步要求，而对后两步的解答则需要结合文献述评深入挖掘。顺着这一思路，研究者发现已有沟通策略的研究存在“较少关注小组互动”及“较少关注综合性任务”的不足，明确新研究想要解决的是关于学生在综合小组讨论中的沟通问题，且沟通策略或技巧的使用无国界，具有重要的实践价值。至此，三步骤的公式全部满足，研究者便可确定这是个有价值的实证研究问题。

明确了研究问题后，研究问题的提出应注意精细化，最好以二至四个（视情况而定）相互联系的小问题来呈现，如：

To understand to what extent the teachers' dissatisfaction is based on sufficient empirical evidence as well as to address the research gap we identified in our literature review,

the present study aimed to (a) explore the relationships between the Listening-Reading-Writing Task and the Independent Listening Task and (b) to examine the language competences assessed by the Listening-Reading-Writing Task. Specifically, we ask four research questions: (a) Is there a meaningful relationship between the scores in the Independent Listening Task and Listening-Reading-Writing Task in the testing of Chinese as a first language? (b) Are there any common factors in the competence assessed by the Independent Listening Task and the Listening-Reading-Writing Task? (c) What competence factors are assessed by the Listening-Reading-Writing Task? (d) Which, if any, performance indicators of the Independent Listening Task can significantly predict scores in the Listening-Reading-Writing Task? (Zhu et al., 2016, 页170)

上述研究旨在探讨两个方面的内容：一是综合写作任务与独立聆听任务之间的关系，二是综合写作任务所评估的语言能力。为此，提出四个精细化的小问题，分别是：（a）独立聆听任务与综合写作任务之间是否存在有意义的关系？（b）独立聆听任务与综合写作任务中是否包含了共同的能力因素？（c）综合写作任务中主要评估的能力因素是什么？（d）独立聆听任务中是否存在某些能力因素可以显著预测学生在综合写作任务中的表现？

实验类的研究论文还通常要求写出假设，因为实验结果根据已有文献及对实验因素的控制，在一定程度上是可预期的。而非实验类的研究论文，如果有足够的文献支持，研究者也可通过推理写出各个问题的研究结果的假设，不少刊物有这方面的要求，但无硬性规定。

简要概括，研究者应始终坚持“问题导向”，在现实场境中找到问题，在文献回归中发展问题。表4概述了达成文献述评的各项功能的策略。为了实现整合功能，宜采用文献的搜寻与管理、梳理与关联的策略。评价功能要注意评价应基于特定的内容，不要泛泛而谈；评价态度中肯、客观。最后，为达成提问功能，则强调结合已有研究的空白点或不足，在现实场境中进一步完善问题。

表4：达成文献述评功能的策略

功能	达成策略	实现细节
整合	搜寻与管理文献	泛搜文献：初步筛选适用文献。
		精选文献：追溯过去文献与追踪新近文献。
		使用管理文献数据库的软件：EndNote / Mendeley等。
	梳理与关联文献	整理与融合文献：采用时间、方法、主题等顺序（其内部逐层接近研究问题）。
		梳理需要特别关注的内容：有关研究设计、结论、建议、研究不足等内容。
评价	关联文献的内容：少单篇罗列，多跨篇总结；如有多段述评同一主题，加小标题；如有长段，加主题句。	
	紧密联系内容	针对特定内容：如文献较长，分几个主题，每个主题可跟随一个相应的评论（初学者较易把握）。
提问	中肯适度批评	保持客观、言辞中肯；批判有据、态度谦逊。
	发展研究问题	以已有研究的空白点、不足为基础，关注现实场境提出研究问题。参考三步骤公式（Booth et al., 2008），回归文献发展问题： （1）题目：我正在研究 _____；（2）问题：想要找出的问题是什么（what）/ 为什么（why）/ 如何（how）_____；（3）重要意义：为了帮助读者了解 _____ 理论 / 实践价值。
	精细化研究问题	提出2-4个相互联系的小问题。 如有扎实的文献，考虑加上假设。

三 建构文献述评框架的意义

写好文献述评，对写好整篇文章也有着重要作用，如为研究工具设计、结果讨论提供一定的参考依据。研究者在概述过往研究的设计，如工具、抽样等与评论不足时，往往也能从中得到发展本研究设计的积极启示，为选择其研究的方法或工具提供参考。另一方面，文献述评也包含了对研究结果的可能性的预测。当实际结果与预期相符时，研究者在结果讨论时，可直接关联支持这一结果的相关研究；当实际结果与预期不符时，研究者便可对照文献归纳异同，尝试从与已有结论相异的角度去解读新结果。

建构文献述评“三阶段*两层次”的立体框架，力图发挥以下三个方面的价值：第一，形成“问题导向”意识，提出新的研究问题是文献述评的核心任务之一。从

文献与现实场境中找出问题，再回归文献完善问题，从而确保研究的前沿性以及创新性。

第二，每个阶段分为两个层次，研究者可遵循由浅至深的路径来完成每一阶段的任务。同时，分层还能帮助研究者及时评估写作，检查自己的写作是否在每个阶段都达到了同等层次的功能。Palmatier等（2018）界定了评估文献述评质量的几个指标：深度、严密性、可还原性以及实际效用，Snyder（2019）整理了用于评估文献述评写作的一些指导原则，如“文献述评的组织结构与研究问题连贯吗？”“文献述评汇报的结果是否足够清晰？”“文献述评中是否涵盖了启示未来研究的问题或方向？”等等。上述结论说明，评估需要设置一定的参照对象。而文献述评框架中的每个阶段、每个层次的功能都能发挥这一作用，方便研究者对照了解自己的写作进展及质量水平。

第三，每个层次横跨三个阶段，研究者可连贯从整合到评价再到提问的写作。就现状而言，不少论文在整合与评价阶段中都能达到内层功能，但在提问阶段中则止步于外层的“奠基”功能，即没有清晰地提出研究问题。究其原因，可能是研究者对现状问题的认识欠深刻，或是研究者只重视文献述评中揭示已有研究成果的功能。而本文框架中，述评的阶段与层次的交叉设计则可帮助研究者克服这一困难，提出基于已有述评的实证研究问题。

四 小结

适当的文献述评能为体现实证研究中的问题的前沿性与结果的原创性奠定基础，这是写好文献述评的重要动机。不过，在中文学术环境中，由于研究者存在对“述”和“评”的理解差异，因此文献述评的写作容易陷入“述评失衡”和“述评失焦”的误区。为了切实提高文献述评的质量，本文提出了“三阶段*两层次”的文献述评框架，强调以“问题为导向”，从已有研究中存在的不足、空白点出发，结合现实教育环境中的现状问题，找出并发展新的研究问题。“三阶段”中的“整合”要求先客观全面地总结已有的研究成果，这一点可参考文献的搜寻与管理策略来完成。在此基础上，通过做好对文献中的重点内容的梳理与关联，架构新研究的变量或模型。“评价”阶段要求特别关注文献中的研究不足和空白之处，在肯定已有研究的价值的同时，从中找出研究的新需求和兴趣。这一阶段是“整合”与“提

问”的过渡阶段，写作时需要注意评价与文献主题之间的关联性，保持客观有理、评论有据。“提问”阶段既是整合与评价的出发点，也是前两个阶段的终点目的。它要求先结合相关的现实场境提出具有国际意义的实践问题，后结合文献的整合与评价进一步完善问题。最终提出的研究问题应具有明确的理论与实践价值，并以精细化的小问题的形式来呈现。两个层次中的“奠基”要求揭示已有研究的客观成果及相关不足，概述现实环境中的现状问题，为“创新”奠定基础。而“创新”则直接导向新研究，在引出研究变量、框架或模型的基础上，明确研究的学术需要、动机，最终提出研究要回答的创新问题。

本文提出的文献述评的写作策略，可为研究者提供参考、借鉴。

【责任编辑：雷秋明】

参考文献：

- Cui, J. J. 崔建军 (2014) Lunwen wenxian zongshu de diwei、xiezuo yuanze yu xiezuo fangfa—Yi jingjixue zhuanke lunwen xiezuo wei li 论文文献综述的地位、写作原则与写作方法——以经济学专业论文写作为例 (Status, writing principles and method of literature review: Taking the paper writing of economics major as example), Tangdu Xuekan 《唐都学刊》 (Tangdu Journal), 30 (5), 117-121.
- Li, X. Y. 李泉鹰 (2011) Wenxian zongshu: Xueshu chuangxin de jishi 文献综述：学术创新的基石 (Literature review: A base of academic innovation), Xuwei Yu Yanjiusheng Jiaoyu 《学位与研究生教育》 (Academic Degrees & Graduate Education), 9, 38-41.
- Barkaoui, K., Brooks, L., Swain, M., & Lapkin, S. (2012). Test-takers' strategic behaviors in independent and integrated speaking tasks. *Applied Linguistics*, 34(3), 304-324.
- Booth, W. C., Colomb, G. G., & Williams, J. M. (2008). *The craft of research* (3rd ed.). The University of Chicago Press.
- Brown, A., Iwashita, N., & McNamara, T. (2005). *An examination of rater orientations and test-taker performance on English-for-academic-purposes speaking tasks*. Educational Testing Service.
- Census and Statistics Department. (2016). *Snapshot of the Hong Kong Population*. HKSAR.

- Cohen, A. D. (2003). The learner's side of foreign language learning: Where do styles, strategies, and tasks meet? *International Review of Applied Linguistics in Language Teaching*, 41(4), 279-291.
- Cohen, A. D. (2010). Focus on the language learner: Styles, strategies and motivation. In N. Schmitt (Ed.), *An introduction to applied linguistics* (2nd ed., pp. 161-178). Hodder Education.
- Cooper, H. M. (1998). *Synthesizing research: A guide for literature reviews* (3rd ed.). SAGE Publications.
- Cumming, A., Kantor, R., Baba, K., Erdosy, U., Eouanzoui, K., & James, M. (2005). Differences in written discourse in independent and integrated prototype tasks for next generation TOEFL. *Assessing Writing*, 10(1), 5-43.
- Douglas, D. (1997). *Testing speaking ability in academic contexts: Theoretical considerations*. Educational Testing Service.
- Faerch, C., & Kasper, G. (1984). Two ways of defining communication strategies. *Language Learning*, 34(1), 45-63.
- Fink, A. (2014). *Conducting research literature reviews: From the internet to the paper* (4th ed.). SAGE Publications.
- Gebril, A. (2006). *Independent and integrated academic writing tasks: A study in generalizability and test method* [Unpublished doctoral dissertation]. The University of Iowa.
- Gebril, A. (2009). Score generalizability of academic writing tasks: Does one test method fit it all? *Language Testing*, 26(4), 507-531.
- Gebril, A. (2010). Bringing reading-to-write and writing-only assessment tasks together: A generalizability analysis. *Assessing Writing*, 15(2), 100-117.
- Guo, L., Crossley, S. A., & McNamara, D. S. (2013). Predicting human judgments of essay quality in both integrated and independent second language writing samples: A comparison study. *Assessing Writing*, 18(3), 218-238.
- Hansen, D. A., & Gerst, J. E. (1967). *On education: Sociological perspective*. Wiley.
- Huang, H. T. D. (2016a). Exploring strategy use in L2 speaking assessment. *System*, 63, 13-27.
- Huang, H. T. D. (2016b). Test-taking strategies in L2 assessment: The test of English for international communication speaking test. *Perceptual and Motor Skills*, 123(1), 64-90.
- Lee, Y. W. (2006). Dependability of scores for a new ESL speaking assessment consisting of integrated and independent tasks. *Language Testing*, 23(2), 131-166.
- Lewkowicz, J. (1994, December). *Writing from sources: Does source material help or hinder students' performance* [Paper presentation]. Annual International Language in Education Conference, Hong Kong.

- Li, D. C. S. (2017). *Multilingual Hong Kong: Languages, literacies and identities*. Springer.
- Machi, L. A., & McEvoy, B. T. (2016). *The literature review: Six steps to success* (3rd ed.). SAGE Publications.
- Nakatani, Y. (2006). Developing an oral communication strategy inventory. *The Modern Language Journal*, 90(2), 151-168.
- National Research Council. (2002). *Scientific research in education*. Committee on Scientific Principles for Education Research. R.J. Shavelson & L. Towne, (Eds.). Center for Education. Division of Behavioral and Social Sciences and Education. The National Academy Press.
- Oxford, R. L. (1990). *Language learning strategies*. Newbury House Publishers.
- Palmatier, R. W., Houston, M. B., & Hulland, J. (2018). Review articles: Purpose, process, and structure. *Journal of the Academy of Marketing Science*, 46(1), 1-5.
- Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research*, 104, 333-339.
- Swain, M., Huang, L.-S., Barkaoui, K., Brooks, L., & Lapkin, S. (2009). *The speaking section of the TOEFL: iBT (SSTiBT): Test-takers' reported strategic behaviors*. Educational Testing Service.
- Watanabe, Y. (2001). *Read-to-write tasks for the assessment of second language academic writing skills: Investigating text features and rater reactions* [Unpublished doctoral dissertation]. University of Hawaii.
- Zhu, X., Li, X., Yu, G., Cheong, C. M., & Liao, X. (2016). Exploring the relationships between independent listening and listening-reading-writing tasks in Chinese language testing: Towards a better understanding of the construct underlying integrated writing tasks. *Language Assessment Quarterly: An International Journal*, 13(3), 167-185.
- Zhu, X., Liao, X., & Cheong, C. M. (2019a). Strategy use in oral communication with competent synthesis and complex interaction. *Journal of Psycholinguistic Research*, 48, 1163-1183.
- Zhu, X., Loh, E. K. Y., Yu, G., Tam, L. C. W., & Liao, X. (2019b). Developing 21st century skills for the first language classroom: Investigating the relationship between Chinese primary students' oral interaction strategy use and their group discussion performance. *Applied Linguistics Review*.