

数字人文视域下证据导向大学英语议论文写作教学的实证研究

袁琪玮

邵阳学院外国语学院，湖南邵阳，422000；

摘要：在数字化信息环境与生成式人工智能进入课堂的背景下，大学英语议论文写作教学面临如何引导学生从“观点表达”走向“证据驱动论证”的现实挑战。基于数字人文视域，本文在一门大学英语写作课程中设计并实施了为期8周的证据导向议论文教学任务链，将多源材料处理、证据表建构、轻量文本分析、语言支架与形成性评价有机融入写作过程，并在“有限允许AI辅助”的教学规则下开展课堂实践。研究采用单组前后测设计，结合写作量表评分、CAF指标、论证要素编码与典型个案分析，对教学成效进行检验。结果表明，学生议论文整体质量显著提升，提升主要体现在证据使用质量与论证质量，论证结构中证据与解释要素显著增加；语言准确性、复杂性与流利性亦得到改善，而篇章连贯提升不显著。研究表明，数字人文视域下的证据导向教学能够有效促进学生证据意识与论证能力的发展，为大学英语议论文写作教学提供了可操作、可评估的课堂路径。

关键词：数字人文；大学英语写作；议论文；证据导向教学

DOI：10.64216/3104-9702.25.07.024

前言

随着大学英语写作教学向“学术化”转型，学生需具备基于证据的论证能力（杨蕊，2024）。然而，当前教学普遍面临三重困境：论证流于观点罗列而缺乏逻辑推进（吉洁等，2015）；数字材料使用陷入“拼贴式写作”且信源不可溯；语言修订止步于表层润色，缺乏连接证据与主张的学术话语支架（Stapleton et al., 2015）。

“数字人文”（DH）为此提供了方法论突破。DH不仅是技术工具，更是一种强调“证据生产”与“可视化解释”的知识生产方式。将DH引入写作教学，核心在于利用轻量分析与语料支架，帮助学生将散乱材料转化为可追溯证据，并建立证据与段落主张的逻辑桥接。

尽管数据驱动学习（DDL）已有较多探索，但如何将数字资源从“语言纠错工具”拓展为“证据建构支架”，并在生成式AI背景下厘清技术辅助的边界，仍缺乏系统的课堂实证。基于此，本研究构建了“数字人文视域下证据导向议论文任务链”，在某高校开展为期8周的教学实验。通过整合多源筛选、证据表建构与可视化写作，并结合前后测、CAF指标及论证要素编码，本研究旨在验证该路径对提升学生论证质量与语言表现的有效性，为二语写作教学提供可复制的闭环方案。

1 文献综述

1.1 数字人文（DH）：从工具技能到证据生产的方

法论

DH不仅是技术工具的集合，更是一种以数字资源处理为基础的知识生产方式。DH pedagogy 研究指出，教学核心应从“软件技能训练”转向培养学生的信息素养、批判性判断与协作能力（Mahony & Pierazzo, 2012）。在二语教学领域，Georgopoulou（2025）指出DH应通过项目式学习嵌入“材料—证据—表达”的链条，而非单纯的技术展示。这为本研究利用数字手段转化散乱材料为可追溯证据提供了理论依据。

1.2 二语议论文写作：论证结构的关键要素与支架需求

基于过程—体裁取向（Process-Genre Approach）与Toulmin论证模型的研究表明，高质量议论文依赖于“主张—证据—推理—反驳”的结构化建构（Badger, R., 2000）。针对中国EFL学习者的研究发现，尽管学生能提出基本主张，但“反驳”与“限定”等高级论证要素显著缺失，且与写作质量呈强相关（Qin&Karabacak, 2010）。这提示教学需超越语言表层，对论证逻辑进行显性支架化。同时，源材料写作（Source-based Writing）研究强调，学术写作的核心在于材料的筛选、整合与规范引用，而这正是当前学生的能力短板（Gebril, 2016）。

1.3 数据驱动学习（DDL）与AI辅助：语言支架的有效性边界

DDL 与语料库辅助写作已被证实能有效提升词汇搭配与语域选择的准确性 (Boulton & Cobb, 2017), 但其贡献主要局限于语言层面的“表达替换”。Yoon&Hirvela (2004) 指出, 若缺乏任务链引导, 学生易陷入数据解读困难。在生成式 AI 背景下, 研究证实 AI 可提升反馈效率, 但也带来证据虚构与思维依赖的风险 (Khampusaen, 2025)。因此, 关键在于界定技术边界, 技术应降低表层语言负担, 但不可替代“证据生产”与“论证建构”的核心认知过程。

综上, 现有研究形成三条支撑链: (1) DH pedagogy 提供了方法论转向; (2) 论证模型揭示了结构短板;

(3) DDL 验证了语言支架的可能。然而, 现有研究仍存在两大缺口: 其一, DH 进入外语写作课堂多停留在理论倡导, 缺乏将“数字资源处理—证据建构—论证推进”整合为可复制任务链的实证; 其二, 生成式 AI 背景下的写作研究多关注反馈效率, 较少与“证据可追溯性”及“论证要素发展”置于同一框架考察。因此, 本研究在 8 周议论文教学中引入证据表、轻量文本分析和可视化与语料支架, 并在“有限允许 AI”的条件下, 结合量表评分、CAF 指标、论证要素编码与个案追踪, 旨在为外语写作课堂提供一套可迁移、可评估的“DH 视域证据导向写作教学”模型。

2 研究设计

2.1 研究问题

本研究拟回答以下三个问题:

(1) 数字人文视域下的证据导向任务链能否显著提升学生议论文写作质量?

(2) 该教学干预对学生 CAF 语言表现有何影响?

(3) 学生论证结构中的关键要素是否发生可观察的变化, 其变化机制为何?

2.2 研究思路

本研究采用单组前后测准实验设计, 结合定量与定性数据进行三角互证。研究对象为湖南某高校 42 名非英语专业大二学生, 其英语水平大致处于大学英语四级

相当区间。研究在为期 8 周的议论文写作必修课程中实施“数字人文视域下证据导向任务链”教学干预, 课程聚焦议论文写作, 议题围绕数字化生活、教育与科技等公共领域。

教学设计遵循“证据生产—证据解释—论证写作—形成性评价—迭代修订”的闭环原则。核心干预要素包括: 第一, 使用“证据表”作为外化支架, 要求学生在起草前填写包含段落主张、证据、解释、来源定位及反方回应的结构化表格, 并随草稿迭代更新; 第二, 嵌入轻量文本分析, 指导学生对多源材料进行关键词与立场识别, 建立证据库; 第三, 提供语料与可视化支架, 通过 DDL 微任务训练论证连接、限定与反驳句式, 并要求学生制作简易图表并撰写解读段, 将其纳入论证链条; 第四, 建立形成性评价链条, 实施“证据表 v1—草稿 1—同伴互评—含图二稿—规范终稿”的迭代流程, 评价焦点始终锁定证据相关性、可追溯性与解释力度。

考虑生成式 AI 的普及, 本研究实行“有限允许 AI 辅助”政策: 允许 AI 用于头脑风暴、语法纠错及表达润色建议, 严禁直接生成整段文本或编造证据、数据。学生需提交 AI 使用声明 (含用途、提示词及修改说明), 并在证据表中明确标注证据来源, 以确保测量聚焦于学生本体写作能力。

2.3 数据收集与分析

研究收集三类数据: 一是写作产品数据, 包括第 1 周前测议论文、第 8 周后测议论文 (限时写作, 题目难度可比) 及项目终稿 (要求含可视化证据与至少 4 条可追溯来源); 二是过程性证据数据, 包括证据表迭代版本、多轮草稿、同伴互评表及学习日志; 三是补充数据, 即课程末期的半结构访谈, 用于解释学生证据意识与 AI 使用体验。

写作质量采用分析性量表评分 (见表 1)。为深入分析论证结构, 基于 Toulmin 模型对文本进行要素编码, 识别并计数主张、理由、证据、解释、反驳与限定等要素, 并细分证据类型 (数据、权威、案例等) 及可追溯性。

表 1 写作评价量表 (分析性量表, 总分 100)

维度 (分值)	评价要点 (操作化)
CAF 语言表现 (30)	准确性: 语法/拼写/用词/搭配错误对理解的影响; 复杂性: 句法结构多样性与恰当性; 流利性: 表达连贯、冗余控制、限时产出充分
篇章结构与连贯 (15)	引言提出可辩论问题并明确立场; 主体段功能清晰、段内逻辑闭合; 衔接手段 (指代/过渡/连接) 有效
论证质量 (25)	主张清晰可辩护; 理由链展开充分; 反驳与限定得当; 推理充分、避免跳论证 (解释证据如何支持主张)
证据使用质量 (20)	证据与主张匹配 (相关性); 证据类型合理 (数据/权威/文本/案例); 证据解释充分; 证据可追溯 (来源可定位)
学术规范与诚信 (10)	文内引用与参考文献基本规范; 改写恰当、避免拼贴; 无编造来源/数据

注：各维度可采用等级评分（如0-4或0-5等级），等级描述建议置于附录。正文报告维度得分与总分。

写作评分与论证编码均由两名具备经验的评分者独立完成，评分前进行试评校准，分歧处由第三方裁决，并计算组内相关系数（ICC）或Cohen's κ 系数以确保信度。定量分析采用配对样本t检验（非正态数据采用Wilcoxon符号秩检验）比较前后测在写作总分、CAF指标及论证要素频率上的差异，报告Cohen's d效应量；定性分析通过论证要素分布统计与典型个案追踪（选取证据意识提升显著的学生），揭示外化支架与形成性评价对论证能力发展的作用机制。

研究已获学生知情同意，所有文本与过程材料在分析中均做匿名化处理，数据仅用于学术研究并加密存储。

3 教学干预实施

3.1 教学目标与能力维度

结合课程定位与议论文体裁要求，本干预将学习目标分为四类：一是论证能力目标，要求学生能够围绕争议议题提出可辩护主张，形成“主张—理由—证据—解释/推理—反驳/限定”的论证链条；二是证据能力目标，要求能够从多源数字材料中筛选、分类并整合证据，保证证据与段落主张对应且来源可追溯；三是语言表达目标，要求掌握议论文常用的论证连接、限定、反驳与证据解释表达资源，提高语言准确性、复杂性与流利性；四是规范与诚信目标，要求掌握基本引用与改写规范，避免拼贴式写作，并在“有限允许AI”条件下形成可记录、可审计的写作过程。干预内容分为三个模块：证据建构（W1-3）、表达与论证（W4-6）、规范与发布（W7-8），具体任务链见表2。

表2 8周证据导向议论文任务链（教学干预实施一览）

周次	核心目标	课堂关键活动	DH方法/工具嵌入点	课后产出物（过程证据）	评价与反馈焦点
W1	建立证据导向写作意识	结构诊断；范文拆解；证据表模板导入；前测写作	证据可追溯规则（示例演示）	前测限时议论文；证据表空表	论证缺口诊断（主张/证据/解释）
W2	多源材料筛选与可信度判断	选题聚焦；来源可靠性判断；材料标注/摘要	数字资源检索；来源标注与分类	材料包（≥6条、含不同立场）；证据表v1	证据可用性与可追溯性
W3	材料→论点：形成论点树	关键词/立场线索提取；对比表；论点树（outline）	轻量文本分析（关键词/立场词/对比表）	发现卡+论点树	论点可辩护性与结构合理性
W4	学术表达支架	DDL微任务：搭配/词汇束/连接语；段落改写训练	语料/例句库（教师提供或自建小语料）	语言资源卡；段落改写作业	语言准确性与语域适配
W5	草稿1：证据嵌入与解释	段落“主张—证据—解释—回扣”；反驳段训练；互评	证据表驱动起草（v2）	草稿1（800-1000词）；证据表v2	证据相关性、解释力度、反驳质量
W6	可视化证据与解释性写作	制作1-2个图表/对比表；写图表解读段并嵌入论证	简易可视化→解释段生成	图表+解读段；草稿2（整合图表）	图表—论点一致性与解释充分性
W7	引用改写与学术诚信	引用与改写训练；拼贴风险识别；规范清单；互评2	引用溯源核查；改写自检表	草稿3；参考文献与文内引用；改写自检表	引用规范、可追溯性、非拼贴改写
W8	成果发布与反思	终稿展示；“证据链”口头说明；后测写作；反思	数字展品式呈现（论点树+证据表+图表）	后测限时议论文；终稿（1000-1200词）；学习反思；AI使用声明	前后对比与迁移反思

注：AI仅限构思与语言层面支持；禁止代写与虚构证据/文献；证据需可追溯、可定位。

3.2 关键教学支架与课堂操作要点

为保证“证据导向”不流于口号，干预中设置了三类核心支架。

第一是证据表支架，作为贯穿全程的外化工具，要求学生在起草前对齐“段落主张、证据、来源定位、解释推理、反方回应”五要素。课堂操作上，教师不直接批改整篇草稿，而是优先批改证据表中的两项：证据是否与段落主张直接相关，以及证据是否可追溯（能定位

到具体出处、页码、段落或链接）。

第二是轻量文本分析支架，为降低工具门槛，重点训练三种可迁移动作：立场线索识别（提取支持或反对立场的关键词与理由句）、证据分类（将材料归入数据、权威研究、文本引证、具体案例等类别）、对比表建构（用对比表呈现两种立场的核心主张与证据以服务反驳段写作）。该支架旨在让学生产生“可写的证据”与“可辩的对立面”。

第三是语料/表达支架，在第4周设置DDL微任务，

强调“论证功能—表达手段”匹配。语言资源卡覆盖四类表达：提出立场（如 argue、claim、maintain）、引入证据（如 according to、data suggest）、解释推理（如 this implies、a plausible explanation is）、反驳与限定（如 however、in most cases）。课堂要求学生在草稿修订时显式标注使用了哪些资源卡表达，以及这些表达如何服务于“证据—推理”连接。

3.3 形成性评价与反馈机制

干预采用“过程证据驱动”的形成性评价路径：从证据表初版（材料可用性）到草稿1（证据嵌入），再到互评（证据相关性与解释力度），接着是草稿2（可视化解释）、草稿3（引用改写），最后到终稿（整体论证与规范）。互评采用聚焦式问题，例如：该段主张是否明确？证据是否直接支持该主张且可追溯？作者是否解释了“证据为何支持主张”？是否存在反方观点或需要限定的情况？教师反馈遵循“先证据后语言”的顺序，先解决证据错配与推理断裂，再处理句法与词汇

层面问题，避免学生将修订简化为表层润色。

4 结果与讨论

4.1 写作总体表现的前后差异

为检验数字人文视域下证据导向议论文任务链的教学效果，本研究对前测与后测写作得分进行配对样本 t 检验，结果见表 3。总体而言，学生后测总分显著高于前测（ $M_{pre}=65.38, SD=5.52; M_{post}=71.24, SD=7.50$ ），差异显著， $t(41)=7.53, p<.001$ ，且效应量较大（ $d=1.16$ ）。维度层面，提升主要集中在“证据使用质量”（ $t(41)=5.77, p<.001, d=0.89$ ）与“论证质量”（ $t(41)=3.24, p=.002, d=0.50$ ），同时“CAF 语言表现”（ $t(41)=3.67, p<.001, d=0.57$ ）与“学术规范”（ $t(41)=2.99, p=.005, d=0.46$ ）也出现显著提升。相比之下，“篇章结构与连贯”提升未达显著水平（ $t(41)=1.59, p=.119$ ），提示本轮 8 周干预对宏观谋篇的促进可能弱于对“证据—推理—规范链条”的促进。

表 3 前后测写作表现差异

指标	前测 M(SD)	后测 M(SD)	Mean Diff (后-前)	t(df)	p	Cohen's d
总分	65.38 (5.52)	71.24 (7.50)	5.86	7.53 (41)	<.001	1.16
CAF (30)	19.83 (3.03)	21.08 (4.01)	1.25	3.67 (41)	<.001	0.57
篇章结构与连贯 (15)	10.53 (2.36)	10.90 (2.85)	0.37	1.59 (41)	.119	0.25
论证质量 (25)	16.15 (3.56)	17.35 (4.28)	1.20	3.24 (41)	.002	0.50
证据使用质量 (20)	11.48 (2.60)	13.80 (3.02)	2.33	5.77 (41)	<.001	0.89
学术规范 (10)	7.39 (1.36)	8.10 (1.73)	0.71	2.99 (41)	.005	0.46

注：Cohen's d (配对) = Mean(差值)/SD(差值)。

从效应分布看，证据维度的提升幅度最大，符合本研究“以证据生产支撑论证写作”的干预目标：任务链将数字资源筛选、证据表建构、图表解释段写作与互评修订纳入同一闭环，从而更直接作用于证据相关性、可追溯性与解释力度等关键指标。

4.2 CAF 指标变化

为进一步呈现语言层面的变化，本研究报告 CAF 可计算指标（表 4）。结果显示，学生后测每百词错误

数显著下降（ $M_{pre}=6.53, SD=1.96; M_{post}=5.49, SD=2.65$ ）， $t(41)=-4.57, p<.001$ ，效应量为 $d=-0.71$ （方向性指标，表示准确性提升）。复杂性指标 Clauses per T-unit 亦显著上升（ $t(41)=2.77, p=.008, d=0.43$ ），表明学生在句法展开与论证型表达资源的组织上有所增强。流利性方面，总词数显著增加（ $t(41)=4.65, p<.001, d=0.72$ ），说明在相同时限条件下学生文本产出能力得到提升。

表 4 CAF 可计算指标前后差异

指标	前测 M(SD)	后测 M(SD)	Mean Diff (后-前)	t(df)	p	Cohen's d
准确性：每百词错误数	6.53 (1.96)	5.49 (2.65)	-1.04	-4.57 (41)	<.001	-0.71
复杂性：Clauses per T-unit	1.41 (0.14)	1.48 (0.19)	0.07	2.77 (41)	.008	0.43
流利性：总词数（限时）	363.14 (69.61)	407.74 (99.45)	44.60	4.65 (41)	<.001	0.72

注：错误数下降代表准确性提升，因此 d 为负值属正常（方向性指标）。

结合教学设计可以推测，语言层面的改善主要来自 两类支架：其一，第 4 周语料/搭配驱动（DDL）微任

务强化了学术表达与论证连接手段；其二，“有限允许AI”政策将AI用途限定在构思与语言润色/纠错层面，从而可能在不可替代写作思维加工的前提下，降低表层错误率并提升表达流畅度。

4.3 论证要素分布的变化

为解释“论证质量”与“证据使用质量”得分提升的具体来源，本研究对前后测文本进行论证要素编码（表6）。编码结果显示，后测文本中Evidence与

Explanation/Warrant的出现频率显著增加：Evidence从前测M=4.12提升至后测M=5.55, $t(41)=3.70$, $p<.001$, $d=0.57$ ；Explanation/Warrant从M=4.21提升至M=6.19, $t(41)=6.24$, $p<.001$, $d=0.96$ 。与此同时，Rebuttal与Qualifier也出现显著提升（均 $p<.01$ ），表明学生更倾向于在论证中处理反方观点并对结论有效范围进行限定。相较之下，Reason要素增幅不显著（ $p=.365$ ），说明本干预对“提出理由”的影响有限，而对“用证据支撑并解释证据如何支持主张”的影响更为突出。

表6 前后测论证要素分布变化

要素	前测 平均每篇次数 M(SD)	后测 平均每篇次数 M(SD)	Mean Diff	t(df)	p	Cohen's d
Claim	6.29 (1.35)	6.64 (1.72)	0.36	2.19 (41)	.034	0.34
Reason	8.36 (1.74)	8.55 (2.23)	0.19	0.92 (41)	.365	0.14
Evidence	4.12 (1.93)	5.55 (2.94)	1.43	3.70 (41)	<.001	0.57
Explanation/Warrant	4.21 (2.07)	6.19 (2.80)	1.98	6.24 (41)	<.001	0.96
Rebuttal	0.98 (0.95)	1.48 (1.17)	0.50	3.11 (41)	.003	0.48
Qualifier	1.14 (0.81)	1.83 (1.34)	0.69	3.45 (41)	.001	0.53

注：编码单位为段落；每段至少识别一个段落主张。

这一分布特征与“证据导向任务链”的目标一致：教学将“证据表（Claim - Evidence - Explanation - Source）”作为外化支架，并在互评与教师反馈中持续聚焦“证据相关性与解释力度”，从而促使学生由“理由堆叠”转向“证据—推理闭环”的论证推进。

4.4 典型个案

为进一步展示能力变化的过程机制，本研究选取学生S进行纵向个案分析（选择标准：前测总体处于班级中位附近，过程性材料完整，后测在证据与论证维度提升明显）。以下为示例式呈现方式：

（1）前测段落特征：主张明确但证据与解释不足

[Claim] Social media harms college students' learning.

[Reason] It distracts them and wastes their time.

[Reason] Many students cannot concentrate in class.

（该段缺少可追溯Evidence，且缺乏Explanation将“分心”如何导致“学习受损”进行论证桥接。）

（2）干预中期（证据表片段示例）

•段落主张（Claim）：社交媒体使用时长与学习投入下降相关

•证据（Evidence+Source）：[来源A：调查报告/研究摘要]（给出具体数据/引文位置）

•解释（Explanation）：说明“时长↑→注意力碎

片化→深度学习时间↓”的推理链

•反驳与回应（Rebuttal）：承认“社交媒体也可用于学习资源获取”，但限定其有效条件（Qualifier）

（3）后测段落特征

[Claim] Excessive social media use is associated with lower academic engagement.

[Evidence] A recent survey of college students reported that heavy users spent fewer hours on sustained study compared with light users [Source A].

[Explanation] This gap matters because sustained study time is closely linked to completing complex assignments; fragmented attention makes it harder to maintain a coherent problem-solving process.

[Qualifier/Rebuttal] However, when social media is used purposefully (e.g., following course-related channels), its negative impact may be reduced.

（可见Evidence与Explanation/Warrant要素更完整，并出现Qualifier/Rebuttal，符合表6的总体趋势。）

学生S的变化表明，“证据表—互评反馈—迭代修订”在写作发展中起到了关键的外化支架作用：证据表将材料处理转化为可操作的段落写作接口，互评将评价焦点稳定在“证据相关性与解释力度”，最终推动文本从“观点陈述”转向“证据驱动论证”。

5 讨论与启示

5.1 主要发现

本研究在8周课堂干预中实施“数字人文视域下证据导向议论文任务链”，并通过前后测比较、CAF指标、论证要素编码与个案追踪对教学效果进行检验。结果显示：（1）学生写作总体质量显著提升，效应量达到较大水平；（2）提升主要集中在“证据使用质量”与“论证质量”，并在论证要素层面表现为 Evidence 与 Explanation/Warrant 显著增长，同时 Rebuttal 与 Qualifier 亦显著提升；（3）CAF 层面准确性、复杂性与流利性均出现显著改善；（4）篇章结构与连贯维度提升未达显著。

从理论层面看，这一结果与二语写作中“过程—体裁融合”路径的观点一致：写作能力发展不仅依赖表层语言改进，更依赖可操作的过程支架与体裁修辞知识的显性化。尤其在议论文体裁中，“证据—解释”桥接与反驳/限定等修辞资源是区分文本质量的重要特征。本研究通过证据表、轻量文本分析与图表解释性写作，把“材料处理”转化为可执行的论证行动，从而在结构上强化了学习者从 Evidence 到 Explanation/Warrant 的生成路径，减少了“证据—结论跳跃”的推理断裂。

同时，本研究的提升模式提示：学生在议论文写作中的核心短板并非“缺少理由”，而在于“缺少可追溯证据”以及“缺少解释证据如何支持主张的推理语句”。这一点也解释了为何 Reason 要素变化不显著而 Explanation/Warrant 增长显著——教学并未简单增加“观点与理由”产出，而是通过证据建构与解释训练，使论证链条更完整、更具学术可检验性。

5.2 机制解释

综合过程性材料与个案追踪，本研究认为教学效果主要由三重机制共同推动：

（1）证据外化机制：证据表将“阅读材料”转化为“段落写作接口”

证据表要求学生将段落主张与证据来源定位绑定，并在“解释栏”中预生成推理句，使证据不再停留在材料堆砌，而成为可直接转化为段落论证的写作资源。这种外化工具降低了从材料到文本的转换门槛，使学习者更容易形成段落闭环。

（2）解释性加工机制：轻量分析与可视化迫使学

习者解释证据

轻量文本分析与图表解读段写作的共同作用在于：它们不仅帮助学生“找到证据”，更要求学生“解释证据”。图表化与对比化的呈现要求学习者对数据或材料差异作出解释，从而自然增加 Explanation/Warrant 语句产出，并促进限定与反驳策略的出现。

（3）评价聚焦机制：形成性评价稳定强化“证据相关性与解释力度”

互评量表与教师反馈坚持“先证据后语言”的顺序，持续追问“证据是否直接支持主张、是否可追溯、解释是否充分”。评价焦点的稳定性使学生修订行为从“表层润色”转向“证据错配修复与推理补强”，形成可观察的结构性改写。

5.3 为何篇章连贯未显著

本研究中篇章结构与连贯维度未达显著提升，可能由以下因素造成：第一，8周干预时长有限，而宏观谋篇能力（整体结构设计、段落序列组织、全篇信息推进）通常需要更长周期的反复练习与跨任务迁移；第二，本研究的教学聚焦有意将注意力放在“证据—推理链条”，在课堂时间分配上对宏观谋篇的显性训练相对不足；第三，限时写作任务对整体结构的展开空间有限，可能低估了学生在项目终稿中呈现的宏观组织能力。后续研究可考虑将“全篇论证图（argument map）”纳入核心支架，并在互评中加入“段落序列与全篇推进”的评价项，以增强对篇章层面的促进。

5.4 “有限允许 AI” 条件下的教学启示

生成式 AI 进入写作课堂后，教学的关键不是简单“禁用或放开”，而是建立可操作的边界与审计机制。本研究的有限允许政策将 AI 限定在构思与语言层面支持，并通过 AI 使用声明、证据表来源定位与引用核查，使 AI 难以替代证据生产与推理建构过程。结果中证据与解释维度的显著提升提示：在明确边界与可追溯约束下，AI 可以作为降低表层负担的工具，而不会必然削弱论证能力培养。对外语写作课堂而言，这一设计提供了一种可兼顾现实可行性与学术诚信的路径。

5.5 教学实践启示

基于本研究结果与机制解释，提出以下可迁移原则供外语写作课堂参考：

（1）证据优先原则：先建证据库与证据表，再写

主段落；把“证据可追溯”作为硬标准。

(2) 解释为核心原则：评价与训练要把重点放在“证据如何支持主张”的解释语句生成上。

(3) 任务链闭环原则：材料筛选—证据表—草稿—互评—修订—终稿，形成稳定迭代。

(4) 轻量方法嵌入原则：工具最小化，方法服务写作；用关键词/对比表/简易图表替代高门槛技术。

(5) AI边界与留痕原则：AI用于构思与语言支持，禁止代写与虚构证据；要求使用声明与来源定位，确保过程可审计。

5.6 研究局限与未来研究方向

本研究仍存在若干局限：第一，采用单组前后测设计，缺少对照组，未来可引入平行班级对照以增强因果解释；第二，样本来自单一学校与课程情境，外部推广性仍需跨校验证；第三，干预周期较短，未能检验宏观谋篇能力与写作迁移的长期效果；第四，尽管采用评分与编码信度控制，写作评价仍可能受评分者经验与文本主题影响。未来研究可在更长周期中加入 **argument map** 等宏观支架，并开展延迟后测与跨体裁迁移检验；同时可进一步细化“证据可追溯性/相关性/解释力度”的指标体系，以实现更精确的机制检验。

参考文献

- [1] Badger, R., & White, G. (2000). A process genre approach to teaching writing[J]. *ELT Journal*, 54(2), 153 - 160.
- [2] Boulton, A., & Cobb, T. (2017). Corpus use in language learning: A meta-analysis. [J] *Language Learning*, 67(2), 348 - 393.
- [3] Gebril, A. (2016). Source-based tasks in academic writing assessment[J]. *Journal of English for Academic Purposes*.
- [4] Georgopoulou, M. S. (2025). Approaches to Digital Humanities Pedagogy: A systematic literature review within educational practice[J]. *Digital Scholarship in the Humanities*, 40(1), 1

21-136.

[5] Mahony, S., & Pierazzo, E. (2012). Teaching skills or teaching methodology? In B. D. Hirsch (Ed.), *Digital Humanities Pedagogy: Practices, Principles and Politics*. Open Book Publishers, 215-225.

[6] Polakova, P., et al. (2024). The impact of ChatGPT feedback on the development of students' writing skills[J]. *Cogent Education*, 11(1): 299801.

[7] Qin, J., & Karabacak, E. (2010). The analysis of Toulmin elements in Chinese EFL university argumentative writing[J]. *System*, 38(3), 444 - 456.

[8] Yoon, H., & Hirvela, A. (2004). ESL student attitudes toward corpus use in L2 writing[J]. *Journal of Second Language Writing*, 13(4), 257 - 283.

[9] Khampusaen, D. (2025). The impact of ChatGPT on academic writing skills and students' perceptions[J]. *LEARN Journal*, 18(1), 963 - 988.

[10] Stapleton, P. & Y. Wu. (2015). Assessing the Quality of Arguments in Students' Persuasive Writing: A Case Study Analyzing the Relationship between Surface Structure and Substance[J]. *Journal of English for Academic Purposes*(17): 12-23.

[11] 吉洁, 梁茂成. (2015). 学习者英语议论文中主语生命度研究[J]. *外语电化教学*(3): 52-58.

[12] 杨蕊. (2024) 我国大学生英语议论文写作中的论证元素研究[D]. 吉林大学.

基金项目：2023年湖南省普通高等学校教学改革研究项(HNJG-20231055)和2025年度湖南省社会科学评审委员会一般课题(XSP25YBC50)。