

# ChatGPT 写作反馈的有效性及应用于英语写作教学的可行性研究

唐小毅

深圳职业技术大学商务外语学院

DOI:10.12238/er.v8i5.6105

**摘要:** 随着生成式 AI 的发展, ChatGPT 作为自动写作评估 (AWE) 工具在二语写作教学中的应用前景逐渐受到关注。本研究探讨了使用 ChatGPT 生成的写作反馈的有效性及其在英语写作教学中的可行性。研究通过对广东省某大学商务英语专业大一学生的实验数据分析, 比较了 ChatGPT 写作反馈与教师反馈对学生语言进步的影响, 结果显示两者在提升写作能力方面效果相似。同时, 学生对 ChatGPT 和教师反馈的偏好各有不同, 部分学生更喜欢教师反馈中的人性化互动, 另一些学生则更青睐 ChatGPT 写作反馈的清晰和具体性。本研究的主要启示是 ChatGPT 写作反馈可以融于英语写作课程, 在英语写作教学中教师可以采用 ChatGPT 反馈与教师反馈结合的混合模式, 以兼顾技术优势与人文关怀, 从而有效提升学生的写作水平。

**关键词:** 二语写作 教师反馈 自动写作评估系统 AI ChatGPT 写作反馈 有效性

A study on the effectiveness of ChatGPT writing feedback and its application in English writing classroom

Xiaoyi Tang

School of Foreign Languages and Business, Shenzhen Polytechnic University

**Abstract:** With the development of generative AI, the application of ChatGPT as an Automated Writing Evaluation (AWE) tool in L2 writing instruction is gaining increasing attention. This study examines the effectiveness of writing feedback generated by ChatGPT and its feasibility in English writing instruction. The research analyzed data from first-year students who majored in Business English at a university in an eight-week quasi-experimental design and weekly surveys. The results showed that both types of feedback were similarly effective in improving writing skills. Students' preferences indicated a near-even split for both types of feedback. The study implies that AI-generated feedback can be incorporated into writing evaluation and a blended approach that utilizes the strengths of both forms of feedback is recommended.

**Keywords:** L2 writing, Teacher feedback, Automated writing evaluation, Artificial intelligence, ChatGPT feedback, effectiveness

## 1. 引言

写作反馈在二语习得和写作教学中至关重要, 能够帮助学生深入理解语言规则并提升写作能力。传统的二语写作课堂依赖教师反馈, 但随着信息技术的发展, 自动写作评估系统 (AWE) 逐渐成为重要的教学工具。这些系统能提供即时反馈, 帮助学生改善语法、拼写等微观特征, 从而使教师和学生能更多地关注写作内容和组织结构 (张珊珊、徐锦芬 2019)。近年来, 生成式人工智能 (AI) 如 OpenAI 的 GPT-4, 提升了 AWE 系统的反馈质量和多样性, 带来了新的发展机遇。本研究旨在探讨使用 ChatGPT 作为 AWE 工具生成的写作反馈对大学生二语写作能力的影响, 并与传统教师反馈进行比较, 分析其在英语写作教学中的可行性。通过对

ChatGPT 与教师反馈的效果差异和学生态度的分析, 为 AI 技术在英语写作教学中的应用提供实证依据和实践指导。

## 2. 文献综述

教师反馈在二语写作教学中被广泛认为是最有效的教学手段之一, 许多研究表明, 教师反馈能显著提升学生的写作质量, 尤其在内容发展、文本组织、语言准确性和语法复杂性方面 (李勇、邓红霞 2012)。教师的个性化反馈可以基于对学生写作水平的深刻理解, 提供有针对性的建议, 帮助学生提高写作能力 (Hyland 2003)。此外, 教师反馈的互动性也是其重要优势之一, 学生可通过与教师的互动反馈深化对写作技巧的理解, 反复修改以提升写作水平 (Goldstein 2005)。然而, 教师反馈也存在一些局限性, 特别是在时间消

耗和教师主观性方面。Jiang & Yu (2006) 指出，并非所有教师反馈都能有效提高学生的写作能力，反馈效果受其目的性、互惠性和意义性等多种因素影响。研究还表明，不同教师的反馈差异可能影响学生的理解和改进效果。

为克服教师反馈的局限性，自动写作评估系统 (AWE) 逐渐成为现代教育技术的重要组成部分。AWE 系统能够快速提供语法、拼写和标点等方面的即时反馈，减轻教师负担 (张荔、邓红霞 2015)。例如，iWrite 和批改网等系统广泛应用于教育领域，通过自然语言处理技术自动检测并纠正语言错误，并提供简单的结构和组织建议。然而，AWE 系统也有其局限，尤其是在处理复杂语言现象时，它无法像教师那样理解深层语境和隐含意义 (Attali & Bursteim 2006)。尽管 AWE 能够提升写作的语言形式，但在内容和创造性方面的支持较为不足。

随着生成式 AI 技术的发展，像 ChatGPT 这样的语言生成技术开始作为 AWE 工具应用于写作反馈中。这类 AI 系统不仅能够纠正语法错误，还能够根据上下文生成自然流畅的反馈 (Brown et al. 2020)。生成式 AI 的一大优势是其处理大量数据的能力，能生成类似人类的文本，为学生提供多样化的反馈。然而，生成式 AI 的应用也面临一些挑战，特别是在处理复杂的语义和上下文关系时存在局限，且在创造性建议方面能力有限，可能导致反馈不准确或误导 (Li & Chen 2023)。此外，学生对 AI 反馈的接受度和信任度仍是一个亟待验证的问题。Dai et al. (2023) 在研究中发现，使用 ChatGPT 提供的写作反馈比教师反馈更易理解且更为详细，但在某些方面与教师反馈保持一致。他们还发现，学生通常对 AI 生成的反馈持怀疑态度，尤其当 AI 反馈与教师反馈不一致时，他们更倾向于信任教师的建议。Huawei & Aryadoust (2023) 指出，学生和教师通常对 AWE 评分持负面看法。Roscoe et al. (2017) 的研究发现，学生对 AWE 系统持谨慎但积极的态度。

综上所述，教师反馈和 AI 反馈在二语写作教学中各有优势与不足。未来的研究应探索如何将两者有效结合，以便为学生提供更全面的写作支持。本研究的主要问题包括：

1) ChatGPT 生成的写作反馈在提升大学生英语写作能力方面的有效性，与教师反馈有何异同？

2) 学生对 ChatGPT 和教师反馈的接受度与偏好如何，哪些因素影响这些偏好？

3) 如何有效整合 AI 技术与传统写作教学，以实现最佳教学效果？

### 3. 研究设计

#### 1) 研究对象

本研究的实验对象为广东省某高校的大一学生，时间为

2023 年秋季学期。研究对象分为两个队：第一队包括 6 个班级 220 名学生 (男生 50 人，女生 170 人)，分为控制组 (CG) 108 人和实验组 (EG) 112 人。控制组由教师提供写作反馈，实验组则由 ChatGPT 生成写作反馈。第二队为英语写作选修课的 50 名学生 (男生 16 人，女生 34 人)，他们来自不同专业，在八周内每周接受来自教师和 ChatGPT 的书面反馈，并参与问卷调查。

#### 2) 研究方法

本研究采用八周的纵向混合重复测量准实验设计，探讨两个研究问题。第一个研究问题使用前测和后测设计，评估学生的语言进步。学生在学期第一周进行前测，然后每周完成一篇 250 字以上的写作任务，并收到来自教师或 ChatGPT 的反馈。随后，学生根据反馈进行修改并提交最终稿。在第十周进行后测，前后测的评分由两名经验丰富的英语教师独立完成。研究将比对实验组和控制组在语法错误率、词汇多样性、句子复杂度、内容和逻辑结构、思想表达清晰度等方面的变化，分析 AI 反馈与教师反馈对学生写作能力的不同影响。

第二个研究问题设计中，第二队学生每周完成写作任务，接受教师和 ChatGPT 的双重反馈，并进行偏好调查。通过问卷评估学生对两种反馈的偏好，问卷包括八个五点量表问题，涵盖反馈的满意度、理解性、帮助程度等维度。同时，问卷询问学生如果只能获得一种反馈，偏好的选择，并要求解释理由。

研究使用 Grammarly、Coh-Metrix、NLTK 等在线分析工具，对文本进行定量分析，结果通过 SPSS 28 软件处理。定性数据则通过问卷调查分析。通过这些数据分析，本研究旨在评估 ChatGPT 与教师反馈在语言改进方面的效果，并探讨学生对两种反馈的偏好及其原因，进一步探讨如何将 AI 技术有效融入写作教学，实现最佳学习效果。

### 4. 研究结果

#### 4.1 两种写作反馈的有效性对比

表 1 展示了实验组 (EG) 和控制组 (CG) 在前测 (T1) 和后测 (T2) 中的得分描述统计，以及 Shapiro-Wilk 检验结果。实验组的得分在测试期间提升了 5.848，控制组则提升了 6.86。前测中，实验组的平均得分为 27.522 (SD = 1.301)，控制组为 26.820 (SD = 1.391)。后测中，实验组的平均得分为 33.370 (SD = 1.908)，控制组为 33.680 (SD = 2.066)。Shapiro-Wilk 检验显示，实验组的得分在前测 ( $W = 0.939, p = 0.167$ ) 和后测 ( $W = 0.934, p = 0.137$ ) 均呈正态分布。控制组在前测中得分呈正态分布 ( $W = 0.926, p = 0.070$ )，但后测不呈正态分布 ( $W = 0.833, p < 0.001$ )。

表 2 的 ANOVA 结果显示 p 值为 0.625，大于显著性水

平 $\alpha$  (0.05)，这表明反馈类型 (ChatGPT 生成的反馈或教师反馈) 对学生的后测得分没有显著影响

表3 报告了前测和后测中实验组与控制组的独立样本 t

检验结果，显示两组在前测 ( $p = 0.30$ ) 和后测 ( $p = 0.38$ ) 中的得分差异不显著，均大于 0.05。这表明，ChatGPT 生成的反馈与教师反馈在提升学生写作能力方面效果相似。

表1 不同时间段的写作分数描述性统计数据

测试	组别	样本	平均值	标准差	偏度	峰度	Shapiro-Wilk 统计值	自由度	显著性
T1	EG	108	27.522	1.301	0.287	-0.505	0.939	16	0.167
T1	CG	112	26.820	1.391	-0.882	0.331	0.926	16	0.070
T2	EG	108	33.370	1.908	-0.361	-1.001	0.934	16	0.137
T2	CG	112	33.680	2.066	1.570	2.698	0.833	16	0.000*

\*  $p < 0.001$

表2 ANOVA 检测结果

来源	平方和	自由度	均方	F 值	P 值	显著性水平 ( $\alpha$ )
组间	1.225	1	1.225	0.245	0.625	0.05
组内		118	5.386			
总计		119				

表3 独立样本 t 检验结果

测试	实验组	控制组	均值差异	t 值	p 值	显著性水平 ( $\alpha$ )
T1	77.46	76.12	1.34	1.07	0.30	0.05
T2	87.66	86.31	1.35	0.9	0.38	0.05

#### 4.2 两种写作反馈对学生写作提升的影响变化对比

表4 报告了实验组和控制组在前测和后测各项评价指标的得分变化。实验结果显示，实验组在语法准确性、词汇多样性和句子复杂性方面的进步显著高于控制组，而控制组在内容和逻辑结构以及思想表达清晰度方面的改善优于实验组。具体结果如下：

1) 语法错误率：实验组从前测的 15.3% 降至后测的 8.2%，而控制组从 14.8% 降至 10.7%。这表明 ChatGPT 反馈在减少语法错误方面比教师反馈更为有效。

2) 词汇多样性：实验组的词汇多样性从前测的 4.1 增加至后测的 5.6，控制组则从 4.0 增加至 5.1。ChatGPT 反馈在鼓励学生使用多样化词汇方面有明显优势。

3) 句子结构复杂性：实验组使用复杂句的比例从前测的 18.4% 增加至后测的 27.3%，控制组从 17.9% 增加到 24.8%。这表明 ChatGPT 反馈更能促进学生使用复杂句子结构。

4) 内容和逻辑结构：控制组在内容和逻辑结构上的评分从前测的 3.2 分提高到 4.6 分，而实验组的评分从 3.1 分提高到 4.1 分。这说明教师反馈在帮助学生组织内容、构建连贯的逻辑结构方面更为有效。

5) 思想表达清晰度：控制组的评分从前测的 3.3 分提高到 4.6 分，而实验组从 3.2 分提高到 4.3 分。教师反馈强调思想的连贯性和表达的准确性，帮助学生在这方面取得更大进展。

通过对实验组和控制组的对比，可以看出，虽然两种反馈方式在改善学生语言能力方面的效果相似，但其影响因素各不相同。教师反馈更注重内容组织和思想表达的清晰度，而 ChatGPT 反馈则在语法准确性、词汇多样性和句子复杂性方面表现更佳。这表明两者在提升学生写作能力上各有优势，未来的写作教学应考虑结合 AI 技术与教师反馈，以实现更全面的写作能力提升。

表4 两种反馈方式对写作得分的影响变化

评价指标	组别	T1 平均分 (标准差)	T2 平均分 (标准差)	分数变化
语法错误率	EG	15.3% (2.1%)	8.2% (1.5%)	-7.1%
	CG	14.8% (2.2%)	10.7% (2.0%)	-4.1%
词汇多样性	EG	4.1 (0.5)	5.6 (0.4)	+1.5
	CG	4.0 (0.6)	5.1 (0.5)	+1.1
句子复杂度	EG	18.4% (3.0%)	27.3% (2.5%)	+8.9%

	CG	17.9% (3.2%)	24.8%(2.8%)	+6.9%
内容和逻辑结构	EG	3.1(0.5)	4.1 (0.5)	+1.0
	CG	3.2(0.6)	4.6(0.4)	+1.4
思想表达清晰度	EG	3.2(0.6)	4.3(0.5)	+1.1
	CG	3.3 (0.5)	4.6(0.4)	+1.3

4.3 学生对反馈的接受度和偏好

八周，学生对 ChatGPT 反馈和教师反馈的偏好趋于均等，各占 50%。

表 5 显示，经过八周的调查，学生对 ChatGPT 反馈的满意度和有用性逐渐上升。从最初的 65% 偏好教师反馈，到第

表 5 每周调查问卷的描述性统计

周	样本	满意度 (AI) 平均值	满意度 (教师) 平均值	可理解性 (AI) 平均值	可理解性 (教师) 平均值	有用性 (AI) 平均值	有用性 (教师) 平均值	总体偏好 (AI) %	总体偏好 (教师) %
1	50	3.7	4.1	3.8	4.2	3.6	4.1	35%	65%
2	50	3.8	4.0	3.8	4.1	3.7	4.1	38%	62%
3	50	3.8	4.0	3.9	4.1	3.7	4.0	40%	60%
4	50	3.9	4.1	3.9	4.2	3.8	4.1	42%	58%
5	50	4.0	4.2	4.0	4.3	3.9	4.2	45%	55%
6	50	4.0	4.1	3.9	4.2	3.9	4.1	46%	54%
7	50	4.1	4.2	4.1	4.3	4.0	4.2	48%	52%
8	50	4.2	4.3	4.2	4.4	4.1	4.3	50%	50%

从定性数据来看，学生们的反馈偏好受多种因素影响。许多学生偏爱教师反馈，原因在于“人与人之间的互动和理解”。面对面的交流能提供一种更“吸引人”和“有趣的学习方式”，这是学生认为 ChatGPT 无法胜任的。同时，能够向教师提出问题并得到即时反馈被许多学生视为关键因素。部分学生认为，通过与教师的互动，不仅能提高写作能力，还能锻炼口语技能。一位学生提到：“ChatGPT 的反馈很有帮助，但修订后可能缺乏个性化”，这表明依赖 ChatGPT 反馈可能导致个人写作风格的丧失。

另一方面，喜欢 ChatGPT 反馈的学生则看重其清晰性、可理解性、一致性和具体性，特别是在学术写作风格和词汇方面。一些学生表示，ChatGPT 能提供详细的学术写作错误反馈。如一位学生提到：“ChatGPT 为我提供了具体反馈，包括错误位置和修改建议。”另一位学生说：“ChatGPT 的反馈很准确，帮我替换了不太学术的词汇。”此外，ChatGPT 不受时间或可用性限制，学生可以随时审阅反馈，这一点也受到学生的高度评价：“ChatGPT 能随时在每个句子上提供准确反馈，而教师因为时间有限，反馈较为有限”

部分学生认为两种反馈形式各有优点，难以取舍。例如，一位学生写道：“我更喜欢两者结合。教师反馈在互动中有助于集中精力，而 ChatGPT 则能及时指出问题。”另一位学生表示：“如果让我选择，我希望同时接受教师和 ChatGPT 的反馈，因为可以从两者中学到更多。”

总体来看，尽管学生认为 ChatGPT 反馈在清晰性和有用性方面有明显优势，但他们仍然非常重视教师反馈所带来的个性化指导和深刻见解。综合来看，将 ChatGPT 与教师反馈结合的混合反馈模式，可能更能满足学生的多样化需求，从而有效提升他们的写作水平。

5. 研究启示

本研究探讨了 ChatGPT 写作反馈与教师反馈对大学生英语写作能力的影响。研究表明，尽管两者在提升学生写作能力方面表现出相似效果，但其作用机制和应用价值有所不同，以下为主要启示：

1) ChatGPT 反馈与教师反馈的互补性

研究发现，ChatGPT 反馈和教师反馈各有优势，二者结合能够产生更理想的教学效果。ChatGPT 在语法、词汇和句子结构的即时纠正方面表现出色，特别是在处理大量学生作业时，能够快速提供一致性的反馈，减轻教师工作负担。然而，在高层次写作能力的培养上，特别是涉及思想深度、逻辑推理和文化背景等内容时，教师反馈具有不可替代的作用。因此，结合 AI 反馈的即时性和教师反馈的深度性，有助于学生在多个维度提升写作能力。

2) 对 AI 反馈局限性的深入理解

虽然 ChatGPT 反馈在速度和具体性方面表现突出，但在理解复杂语境和提供个性化反馈上存在局限，特别是在高级语言表达和文化敏感性问题上，可能会误导学生。因此，

教师在教学中应引导学生正确使用 AI 反馈, 帮助学生识别和纠正 AI 的潜在错误建议, 从而提高学生利用 AI 反馈的效率, 避免过度依赖 AI。

### 3) 学生偏好对教学策略的影响

调查显示, 大约一半的学生偏好教师反馈, 另一半倾向于 ChatGPT 反馈。那些喜欢教师反馈的学生认为与教师的面对面互动能够增强情感联系、提高参与度, 并有助于口语能力的提升。而偏好 ChatGPT 反馈的学生则看重其反馈的清晰性、具体性和一致性, 认为这些特征有助于改进写作。

需要注意的是, 本研究限制了学生访问 ChatGPT 以防止其用 AI 完成作业, 因此学生未能向 ChatGPT 提出跟进问题。若允许学生向 AI 提出问题, 可能会增加他们对 AI 反馈的偏好。

学生对反馈的偏好提醒我们, 教学设计应更加灵活和个性化。在实际教学中, 教师应根据学生的反馈偏好选择合适的反馈形式, 或者结合两种反馈形式, 以满足不同学生的需求。此外, 教师还应帮助学生合理利用 AI 反馈, 提高学习效果。

### 4) 未来研究方向

本研究的局限性在于样本规模较小, 且仅限于商务英语专业和写作选修课学生。未来的研究可以扩大样本范围, 涵盖不同专业和年级的学生, 验证研究结果的普适性。同时, 还应探索 AI 反馈在不同类型写作任务中的效果, 如创造性写作、学术写作和商务写作等, 从而更全面地评估 AI 技术在写作教学中的应用潜力。

## 6. 结语

随着 AI 技术的迅速发展和广泛应用, 外语教师对其潜力和未来在语言教育中的影响将变得愈发熟悉和重要。尽管目前的 ChatGPT 尚未专门为自动写作评估 (AWE) 优化, 但已有一些现有的 AWE 程序开始将 GPT 技术整合进来 (如 GrammarlyGO)。未来, 大型语言模型有望被进一步优化和调整, 以便更好地服务于语言学习与教学, 尤其是在写作评估方面。随着公众对这些技术的接触增多, AI 很可能成为教育机构常用的学习管理系统的一部分, 成为教师和学生日常实践中不可或缺的工具。

总之, 随着 AI 技术在语言教育中的渗透, 教师应充分利用其带来的优势, 同时深刻认识到人类教育的独特价值。英语写作教师应努力在 AI 技术与人类教师的反馈之间找到平衡, 创建多元化且有效的反馈模式, 从而更好地提升学生的英语写作水平。

### [参考文献]

[1] Attali, Y., & Burstein, J. Automated essay scoring with e-rater® V. 2. *The Journal of Technology, Learning,*

*and Assessment [J].* 2006,4 (3).

[2] Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., ... & Amodei, D. Language models are few-shot learners [M]. 2020, Curran Associates Inc.

[3] Dai, Y., Gao, Y., & Wang, X. Exploring the effectiveness of ChatGPT in providing corrective feedback for undergraduate writing. *Computer Assisted Language Learning [J].* 2023(4): 467-488.

[4] Davis, S., Smith, P., & Zhang, Y. Trust and skepticism: Students' perceptions of AI-generated feedback in writing courses. *Journal of Academic Writing [J].* 2023(1): 1-19.

[5] Goldstein, L. M. Teacher written commentary in second language writing classrooms [M]. 2005, University of Michigan Press.

[6] Hyland, K. Second language writing [M]. 2003, Cambridge University Press.

[7] Jiang, L., & Yu, S. . Exploring the effects of different types of teacher feedback on EFL student writing performance. *Language Teaching Research [J].* 2021(6): 913-931.

[8] Huawei, W., & Aryadoust, V. A review of automated writing evaluation systems: How do they compare with human raters? *Language Testing in Asia [J].* 2023(1): 19-34.

[9] Li, M., & Chen, Q. Challenges in the use of AI-based writing evaluation tools in EFL contexts. *Journal of Educational Computing Research [J].* 2023(3):789-810.

[10] Roscoe, R. D., Wilson, J., Johnson, A. C., & Maynard, K. The effects of automated writing evaluation on learner attitudes and revision behavior: A mixed-methods study. *Journal of Educational Computing Research [J].* 2017(6): 884-906.

[11]李勇, 邓红霞. 教师书面修正性反馈对英语写作复杂度的影响 [J].*外语教学*, 2012, (2): 55-62

[12]张荔, 盛越. 自动作文评阅系统反馈效果个案研究 [J]. *外语电话教学*, 2015, (5): 38-44

[13]张姗姗, 徐锦芬. ZPD 视角下在线自动反馈对英语不同水平学习者写作的影响 [J].*外语与外语教学*, 2019, (5): 30-39.

### 作者简介:

唐小毅, 1974-, 女, 汉族, 广西人, 硕士, 讲师, 主要研究方向: 应用语言学, 英语教学。

### 基金项目:

2023 深圳职业技术大学教育教学研究重点项目“课堂话语分析框架下英语教师学科教学知识研究”