

# 生成式人工智能辅助大学生英语写作效能研究

刘秀梅, 董 静, 田静莹

中国矿业大学(北京)文法学院, 北京

收稿日期: 2025年5月21日; 录用日期: 2025年6月19日; 发布日期: 2025年6月27日

## 摘 要

本研究旨在探讨生成式人工智能辅助大学生英语写作的整体效能及具体表现。研究采用实验法, 共15名大学生志愿者参加为期8周的雅思写作任务2实验。实验中学生完成作文一稿后, ChatGPT4.0为学生提供评改意见, 在此基础上, 学生完成二稿。通过对比实验前测和后测结果, 本研究发现大多数学习者通过AI辅助训练, 写作整体得分都有提升; 在雅思写作任务2的四个评分维度中, 学生在衔接、词汇以及语法三个方面提升最为明显。研究肯定了生成式人工智能辅助大学生英语写作的有效性, 提出进一步探索教师与AI协作开展大学生英语写作评改模式的可能性。

## 关键词

生成式人工智能, 英语写作, 效能研究

# Efficacy of Generative AI in Assisting College Students with English Writing

Xiumei Liu, Jing Dong, Jingying Tian

School of Law and Humanities, China University of Mining and Technology, Beijing

Received: May 21<sup>st</sup>, 2025; accepted: Jun. 19<sup>th</sup>, 2025; published: Jun. 27<sup>th</sup>, 2025

## Abstract

The current study aims to investigate the overall efficacy and specific impacts of generative AI in assisting college students with English writing. An experimental method was employed, involving 15 volunteer university students who participated in an 8-week IELTS Writing Task 2 experiment. After completing their first drafts, students received feedback and revision suggestions from

ChatGPT-4.0, based on which they produced their second drafts. By comparing pre-test and post-test results, the study found that most learners demonstrated improved overall writing scores through AI-assisted training. Among the four IELTS Writing Task 2 assessment criteria, students showed the most significant progress in cohesion and coherence, lexical resource, and grammatical range & accuracy. The findings affirm the effectiveness of generative AI in supporting English writing development and suggest further exploration into collaborative teacher-AI feedback models for university-level writing instruction.

## Keywords

Generative AI, English Writing, Efficacy Study

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

以 ChatGPT 为代表的生成式人工智能以其强大的自然语言处理功能引起社会各界的广泛关注, 促进教育领域的变革发展。“科技与教育的系统性融合是未来教育生态的关键特征”[1]。随着教育数字化转型的推进, 人们对人工智能的态度从“拒之门外”逐渐转为“欣然拥抱”。外语教学领域的专家学者关注数字化教学, 探索人工智能在教学领域的应用潜能, 创新语言学习模式。然而, 现有研究多从理论层面探讨人工智能辅助英语写作的可能性[2], 对人工智能辅助英语写作的实际效能则关注不足。基于此, 本研究通过教学实验, 探索生成式人工智能辅助大学生英语写作的实际效能, 以期为生成式人工智能融入英语写作教学实践提供参考。

## 2. 研究背景

2022 年, OpenAI 公司推出一款生成式智能聊天机器人 ChatGPT, 以其语言理解、文本生成、对话生成等强大的自然语言处理功能为特点, 带给人们全新的智能化体验。Chat GPT 的诞生引起了教育界的广泛关注, 不少学者开始探索生成式人工智能在二语教学领域的应用潜力。生成式人工智能具有强大的自然语言处理功能, 可以为二语学习者提供语言学习资料、帮助学习者练习口语对话[3], 提高语言技能; 还可应用于智能辅助写作, 完善写作大纲, 通过提供个性化反馈, 帮助学习者改善语言形式[4]。ChatGPT 还可辅助教师教学工作, 提供教学资源支持, 辅助二语教师进行自动评估和反馈[3]。

目前, 已有研究从学生认知、教师角色和反馈特点等角度探索生成式人工智能在二语写作教学中的应用。徐林林等(2024) [5]运用质性主题分析法, 探索了生成式人工智能辅助学术英语写作的学习者认知及行为, 学习者认为人工智能主要从提升写作质量和写作效率两个方面辅助学术英语写作, 直接使用、参考加工、迁移应用为学习者主要的行为模式。陈莱和吕明臣(2024) [2]从活动理论视角出发, 分别分析了写作前阶段、写作阶段、评阅阶段及反思阶段教师的角色变化, 提出教师要在 ChatGPT 参与的英语写作教学中充当人工智能技术的使用者、培训者和引导者。通过对 ChatGPT 与教师对学习者的英语作文的反馈比较, 魏爽和李璐遥(2023) [6]提出 ChatGPT 与教师反馈各具特点, ChatGPT 能提供即时、全面的反馈, 弥补教师反馈在时间上的限制, ChatGPT 与教师反馈的结合有助于提升作文批改质量。

然而, 将生成式人工智能应用于英语写作教学仍存在一些争议。在写作中使用人工智能, 有造成学术剽窃的风险[7], 同时还存在一些道德和隐私方面的担忧[8]。Wu (2023) [9]提出 ChatGPT 的快速信息处

理可能会削弱学生的批判性思维能力、导致认知负荷的增加。Barret & Pack (2023) [10]采用问卷的方法比较学生和教师对写作过程中使用生成式人工智能的看法, 提出透明性在生成式人工智能的使用中至关重要, 并表示教育机构需要制定明确的政策以指导生成式人工智能在教育领域的使用。Dang & Wang (2024) [11]采用主题分析方法, 揭示了美国大学在生成式人工智能使用方面的立场, 其中绝大部分大学提到了学术诚信、知识产权、偏见和刻板印象问题。Bašić 等(2023) [12]进行对照实验, 测试 ChatGPT 辅助写作能否提高学生论文写作表现, 结果显示, 对照组(未使用 ChatGPT)得分略高于实验组(有 ChatGPT 辅助), ChatGPT 辅助写作并不一定能提高学生写作质量。

综上可见, 生成式人工智能具有应用到英语写作教学中的潜力, 但也存在一些担忧和隐患。目前的研究多以理论研究为主, 生成式人工智能辅助英语写作的实际效能如何, 能否提高学习者写作能力, 还需更多实证研究的验证。

3. 研究问题

本研究探究生成式人工智能辅助大学生英语写作的效能, 具体回答两个问题:

- (1) 生成式人工智能辅助大学生英语写作的总体效能如何?
- (2) 生成式人工智能可在哪些方面提高大学生英语写作技能?

4. 研究设计

4.1. 研究对象

本研究在北京某“211”高校招募志愿者, 开展为期 8 周的雅思写作教学实验。实验要求测评者之前未参加过雅思考试或相关培训, 以控制变量, 且保证测评者性别比例大致均衡, 学科背景分布均匀, 以减小无关因素对实验结果的影响。最终筛选到测评者人数为 15 人, 15 名测评者均自愿参与且知晓实验相关情况。实验参与者具体信息见表 1, 其中男生 6 人(40%), 女生 9 人(60%); 本科生 10 人(67%), 硕士研究生 5 人(33%), 学科背景为人文社会学科 7 人(47%)和理工科 8 人(53%), 英语水平介于大学英语四级和六级之间。研究团队对所有测评者编号(S1~S15), 保护实验参与者隐私。

Table 1. Basic information of the experiment participants  
表 1. 实验参与者基本信息结构

| 变量   | 类别          | 人数 | 百分比(%) |
|------|-------------|----|--------|
| 性别   | 男           | 6  | 40     |
|      | 女           | 9  | 60     |
| 学段   | 本科生         | 10 | 67     |
|      | 研究生         | 5  | 33     |
| 英语水平 | 未通过 CET4 考试 | 8  | 54     |
|      | CET4 水平     | 2  | 13     |
|      | CET6 水平     | 5  | 33     |
| 学科背景 | 人文社会学科(英语)  | 7  | 47     |
|      | 理工科         | 8  | 53     |

## 4.2. 研究方法

本研究采用实验法,通过比较前后测结果确定实验参与者英语写作能力是否提高。测评者共完成 8 个雅思写作任务 2,雅思写作任务 2 是雅思写作考试的第二个模块,考察学习者的语言运用能力、逻辑思维能力和论证能力。测评者需要在 40 分钟内完成一篇不少于 250 字的议论文,主题涉及社会、文化、科技、教育等。雅思写作 2 从任务回应、连贯与衔接程度、词汇资源与语法和准确性四个维度进行评分。实验中每个写作任务采用“初稿-修改-二稿”的循环模式。测评者在完成每个任务初稿后,在 ChatGPT 的修改建议下完成二稿。研究团队邀请两位雅思写作教学经验丰富的教师(分别编码为 T1、T2)评定测评者实验前测和后测的写作分数,通过比照结果来确定测评者写作能力变化。

## 4.3. 数据收集与分析

写作实验前由两位教师(T1、T2)对 15 名测评者(S1~S15)进行雅思写作任务 2 相关培训,介绍雅思写作任务 2 的基本要求,并进行 Chat GPT 使用培训,确保每位测评者顺利开展雅思写作任务 2 并使用生成式人工智能获取作文评分及具体修改意见。

进入写作实验阶段后,测评者每周需在指定时间的 40 分钟内完成一篇雅思写作任务 2。完成初稿后,为减少其他因素的影响,测评者使用统一提示词“Please grade my writing according to the rating criterion of IELTS Writing Task 2, give global score as well as the separate grades for Task Response, Coherence and Cohesion, Lexical Resource, and Grammatical range and Accuracy.”来获取 ChatGPT 的评改建议。同时,因不同测评者作文中的错误点不同,研究团队鼓励测评者根据作文一稿的实际情况,就进一步修改建议向 ChatGPT 追问,以获得更有针对性的个性化建议。在此基础上,测评者修改作文形成二稿。

实验累计 8 周,研究人员共计收集到 15 名测评者 120 份作文以及由两位教师对 120 份作文的成绩评定。研究人员对两位教师第一次和第八次的评分数据进行定量分析,使用 SPSS 计算前测、后测的整体及各维度打分增量、提分均值、提分方差等。

## 4.4. 生成式 AI 辅助大学生英语写作的总体效能

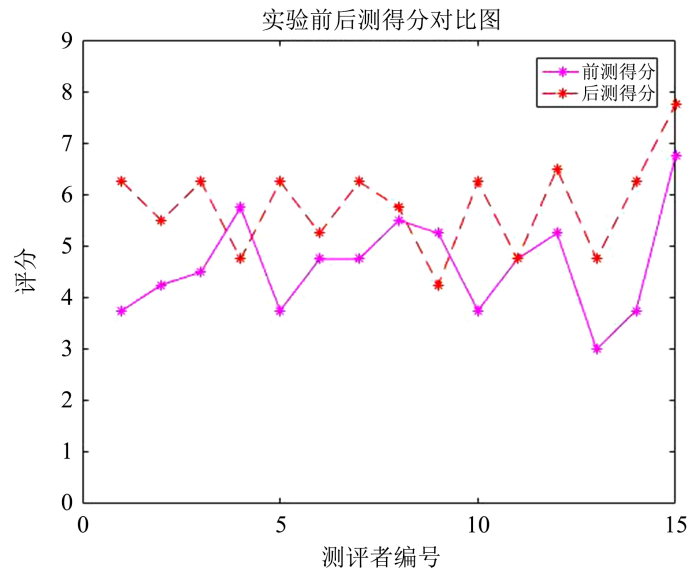
通过比较 15 名实验参与者前测与后测的得分数据以及总分的得分均值来验证生成式人工智能辅助大学生英语写作的总体效能,具体结果见图 1 和表 2。如图 1 所示,横坐标代表测评者编号 S1~S15,纵坐标代表评分数值 0~9。为减小教师评分主观倾向性的影响,统计时取 T1、T2 两位教师的评分均值作为测评者总分最终得分数据。图 1 中共有两条曲线,实线代表前测得分,即两位教师对测评者第一次测试的总分评分均值;虚线代表后测得分,即两位教师对测评者第八次测试的总分评分均值。曲线上的点分别代表每一位测评者前测及后测的总分得分具体数值。观察图 1 可知,除个别点位外,虚线大致分布于实线上方。说明绝大部分测评者后测得分高于前测得分,写作总成绩提高。分别将 15 名测评者后测得分与前测得分作差,获得 15 名测评者的整体得分增量(表 2)。由表 2 数据可知,除两位测评者外,其余测评者整体得分增量均为正值。计算 15 名测评者整体得分增量均值为 1.150,即每名测评者后测比前测总分平均提升 1.150 分。这说明大多数测评者在经过八次的生成式人工智能辅助英语写作训练后,写作整体得分提高,英语写作总体水平提升。极少量测评者后测写作得分并未提高,这与个人学习方法偏好、知识接受速度等有关,并不影响实验总体结果。

## 4.5. 生成式人工智能辅助大学生英语写作的具体表现

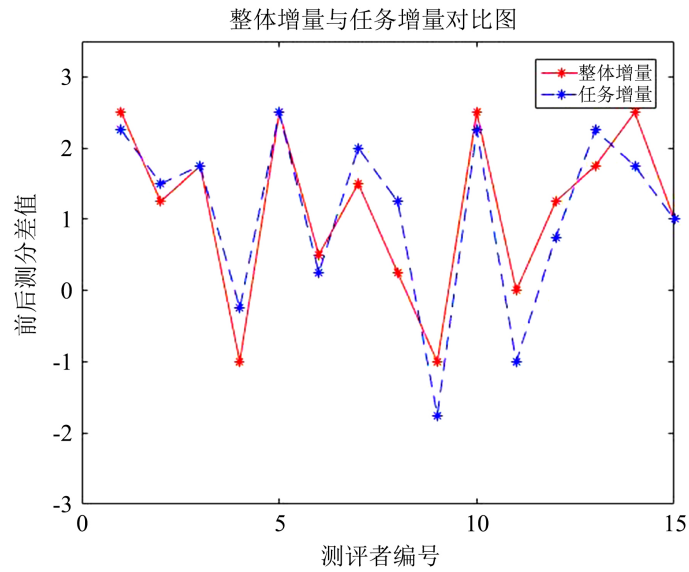
通过比较 15 名测评者整体增量与各维度增量的一致程度,以及总分与各维度提分均值来探讨生成式人工智能辅助大学生英语写作的具体表现,详细结果见图 2、图 3、图 4、图 5 和表 3。

**Table 2.** Participants' overall score increment  
**表 2.** 测评者整体得分增量表

| 测评者编号 | 整体得分增量 | 测评者编号 | 整体得分增量 | 测评者编号 | 整体得分增量 |
|-------|--------|-------|--------|-------|--------|
| S1    | 2.50   | S6    | 0.50   | S11   | 0      |
| S2    | 1.25   | S7    | 1.50   | S12   | 1.25   |
| S3    | 1.75   | S8    | 0.25   | S13   | 1.75   |
| S4    | -1     | S9    | -1     | S14   | 2.50   |
| S5    | 2.50   | S10   | 2.5    | S15   | 1      |



**Figure 1.** Pre-experiment vs. post-experiment score comparison  
**图 1.** 实验前后测得分对比图

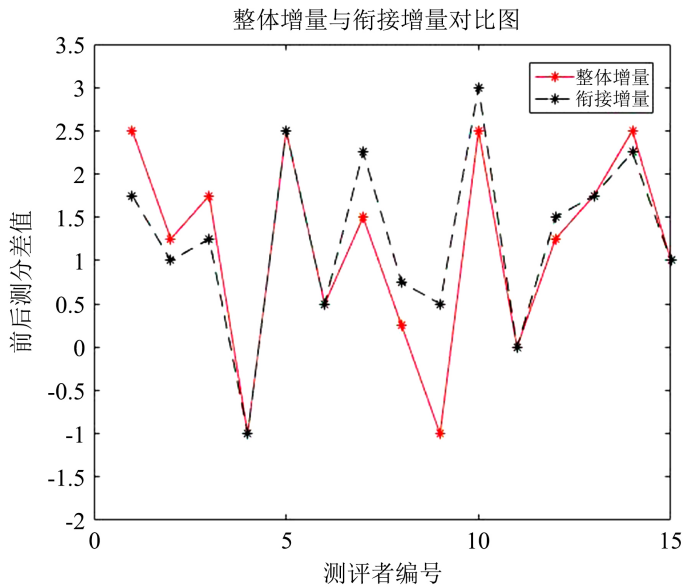


**Figure 2.** Overall increment vs. task response increment  
**图 2.** 整体增量与任务增量对比图

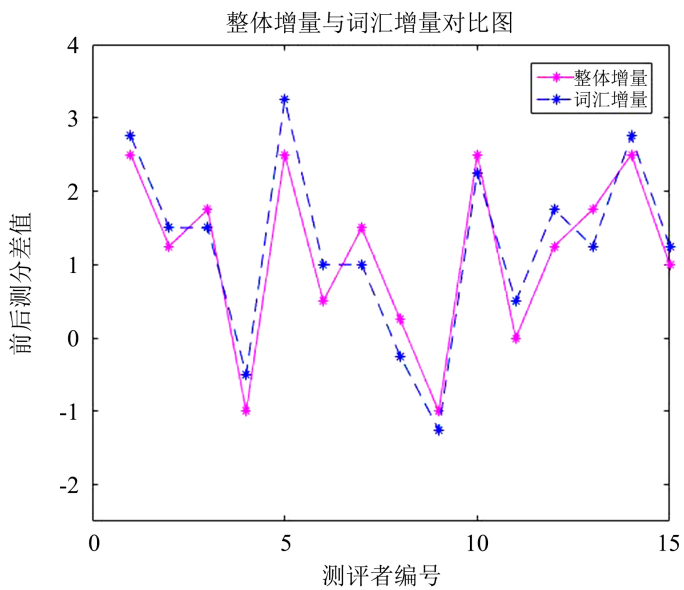
**Table 3.** Average and variance of score gains (total & by dimension)  
**表 3.** 总分及各维度提分均值与方差

| 项目   | GR    | TR    | CC    | LR    | GRA   |
|------|-------|-------|-------|-------|-------|
| 提分均值 | 1.150 | 1.100 | 1.267 | 1.250 | 1.250 |
| 提分方差 | 1.239 | 1.363 | 1.124 | 1.430 | 1.436 |

注：GR 为整体得分，TR 为任务响应得分，CC 为衔接得分，LR 为词汇得分，GRA 为语法得分。



**Figure 3.** Overall increment vs. coherence increment  
**图 3.** 整体增量与衔接增量对比图



**Figure 4.** Overall increment vs. vocabulary increment  
**图 4.** 整体增量与词汇增量对比图

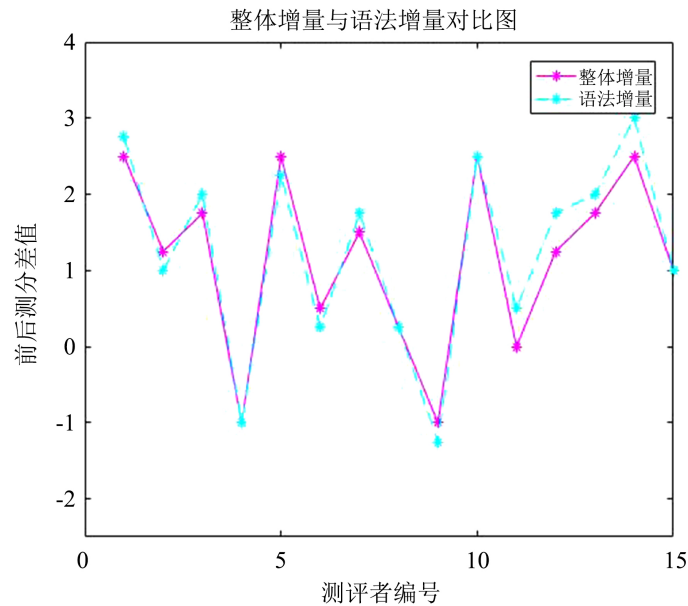


Figure 5. Overall increment vs. grammar increment  
图 5. 整体增量与语法增量对比图

分别将 15 名测评者前测任务得分与后测任务得分作差,即第八次测试两位教师任务维度评分均值减第一次测试两位教师任务维度评分均值,获得 15 名测评者任务得分增量数据,并绘制整体增量与任务增量对比图(见图 2)。图 2 中,虚线代表任务增量,实线代表整体增量,曲线上的每个点分别表示 15 名测评者任务增量与整体增量的具体数值。用同样的方法获得 15 名测评者衔接增量、词汇增量、语法增量,可得图 3、图 4、图 5。分析图 2 至图 5 的曲线走向可知,测评者各维度得分增量与整体增量变化趋势高度一致。计算 15 名测评者任务、衔接、词汇、语法增量均值可得表 3。表 3 表明,学习者在衔接、词汇、语法维度提分均值最高,分别为 1.267、1.250、1.250。雅思写作任务响应,衔接、词汇与语法维度满分分别为 9 分,提高 1.2 分以上可判定为显著进步。

由此可知,通过八次生成式人工智能辅助英语写作训练,测评者在衔接、词汇及语法维度提升最大,任务维度提升最小。这一方面是由于生成式人工智能在评改过程中能够提供丰富的连接词和过渡语、更准确的词汇选择、实时纠正语法错误,帮助测试者优化文章结构、积累词汇搭配、理解语法规则;另一方面是因为学习者本身在雅思写作 2 的任务响应维度面临较小困难。

## 5. 结语

本研究采用教学实验方法探讨生成式人工智能辅助大学生英语写作的效能。总体而言,通过比较测评者前测与后测得分数据及总分的提分均值,研究发现生成式人工智能辅助大学生英语写作可以提高学习者总体写作表现;进一步分析测评者整体增量与各维度增量一致程度及总分与各维度提分均值可知,生成式人工智能辅助英语写作帮助学习者在衔接、词汇和语法维度提升显著。本研究肯定了生成式人工智能提供写作反馈的有效性,这一发现与前人研究结果一致[13]。

鉴于生成式人工智能辅助英语写作教学的潜能,二语教师可与 AI 协作,充分利用 AI 工具辅助学生作文评改,提供补充性反馈;二语学习者可以通过 AI 获得个性化、即时反馈,帮助提高语法准确性、拓展词汇量,有效修改文章。同时,也要意识到生成式人工智能辅助大学生英语写作的一些潜在弊端,如造成学习者对 AI 工具的过度依赖、限制学习者创造性思维等[14]。

## 参考文献

- [1] 张绒. 生成式人工智能技术对教育领域的影响——关于 ChatGPT 的专访[J]. 电化教育研究, 2023, 44(2): 5-14.
- [2] 陈莱, 吕明臣. ChatGPT 环境下的大学英语写作教学[J]. 当代外语研究, 2024, 24(1): 161-168.
- [3] 胡壮麟. ChatGPT 谈外语教学[J]. 中国外语, 2023, 20(3): 1+12-15.
- [4] 郭茜, 冯瑞玲, 华远方. ChatGPT 在英语学术论文写作与教学中的应用及潜在问题[J]. 外语电化教学, 2023(2): 18-23+107.
- [5] 徐林林, 胡杰辉, 苏扬. 人工智能辅助学术英语写作的学习者认知及行为研究[J]. 外语界, 2024(3): 51-58.
- [6] 魏爽, 李璐遥. 人工智能辅助二语写作反馈研究——以 ChatGPT 为例[J]. 中国外语, 2023, 20(3): 33-40.
- [7] Fyfe, P. (2022) How to Cheat on Your Final Paper: Assigning AI for Student Writing. *AI & Society*, **38**, 1395-1405. <https://doi.org/10.1007/s00146-022-01397-z>
- [8] Kumar, R. (2023) Faculty Members' Use of Artificial Intelligence to Grade Student Papers: A Case of Implications. *International Journal for Educational Integrity*, **19**, Article No. 9. <https://doi.org/10.1007/s40979-023-00130-7>
- [9] Wu, Y. (2023) Integrating Generative AI in Education: How ChatGPT Brings Challenges for Future Learning and Teaching. *Journal of Advanced Research in Education*, **2**, 6-10. <https://doi.org/10.56397/jare.2023.07.02>
- [10] Barrett, A. and Pack, A. (2023) Not Quite Eye to A.I.: Student and Teacher Perspectives on the Use of Generative Artificial Intelligence in the Writing Process. *International Journal of Educational Technology in Higher Education*, **20**, Article No. 59. <https://doi.org/10.1186/s41239-023-00427-0>
- [11] Dang, A. and Wang, H. (2024) Ethical Use of Generative AI for Writing Practices: Addressing Linguistically Diverse Students in U.S. Universities' AI Statements. *Journal of Second Language Writing*, **66**, article 101157. <https://doi.org/10.1016/j.jslw.2024.101157>
- [12] Bašić, Ž., Banovac, A., Kružić, I. and Jerković, I. (2023) Better by You, Better Than Me? ChatGPT-3 as Writing Assistance in Students' Essays. <https://arxiv.org/abs/2302.04536>
- [13] Dai, W., Lin, J., Jin, H., Li, T., Tsai, Y., Gašević, D., et al. (2023). Can Large Language Models Provide Feedback to Students? A Case Study on ChatGPT. 2023 *IEEE International Conference on Advanced Learning Technologies*, Orem, 10-13 July 2023, 323-325. <https://doi.org/10.1109/icalt58122.2023.00100>
- [14] Jiang, J., Vetter, M.A. and Lucia, B. (2024) Toward a "More-Than-Digital" AI Literacy: Reimagining Agency and Authorship in the Postdigital Era with ChatGPT. *Postdigital Science and Education*, **6**, 922-939. <https://doi.org/10.1007/s42438-024-00477-1>