

Effects of an AI-supported approach to peer feedback on university EFL students' feedback quality and writing

論文・文献講読

22266505 岩井 政樹

金子研究室 B4

2025-05-23

Outline

1. 概要	2	5.1 レビューの質への効果	20
1.1 実験と結果の概要	3	5.2 ライティング能力への効果	21
2. 研究背景	4	6. 私見	22
2.1 ELF Writing の授業が抱える課題 ..	5	6.1 疑問点	23
2.2 peer review とは?	7	6.2 応用可能性	24
3. 研究目的	9	7. 参考文献	25
3.1 研究目的	10	7.1 参考文献一覧	26
4. 実験手法	11		
4.1 実験環境	12		
4.2 レビュー方法	13		
4.3 peer review の評価基準	14		
4.4 実験の流れ	15		
4.5 分析方法と使用データ	18		
5. 結果	19		

1. 概要

1.1 実験と結果の概要

- ELF Writing（外国語としての英語）の授業で AI Chat Bot「Eva」をオンラインレビューシステムに組み込んだ
- 他の受講者が作成した小論文へのレビューに対して **Eva からのフィードバックを受けた群**はそうでない群と比べて
 - より質の高いレビューができるようになった
 - ライティング能力も向上した [1]

用語の注意

人間からのレスポンスを**レビュー**

AI からのレスポンスを**フィードバック**と表記します

[1]Kai Guo, Mengru Pan, Yuanke Li and Chun Lai. Effects of an ai-supported approach to peer feedback on university efl students' feedback quality and writing ability. The Internet and Higher Education, Vol. 63, p. 100962, 2024.

2. 研究背景

事実

レビューはライティング能力向上に（間接に）効果あり[2]

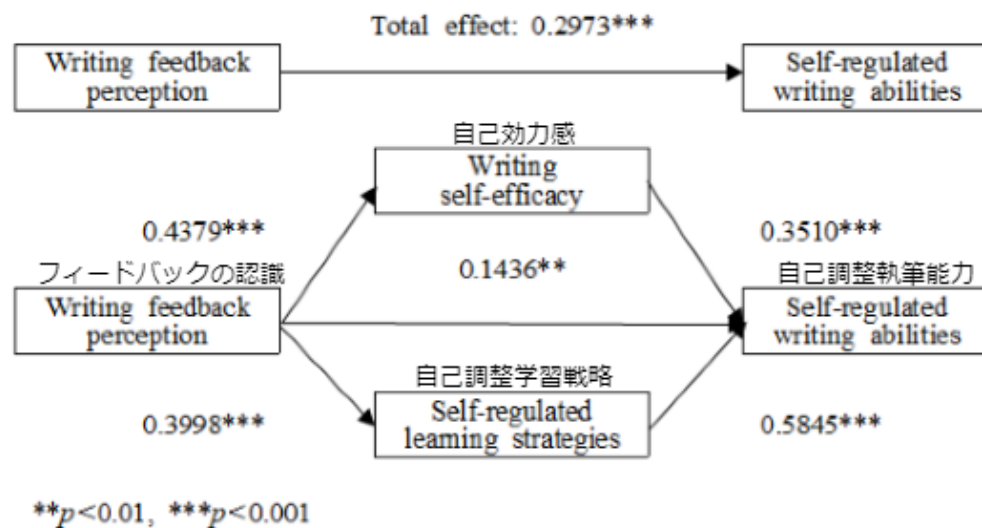


Figure 1: フィードバックが執筆能力に影響を与える経路 出典: Hwang S. (2025)[2] より引用改変

[2]Hwang S. Unpacking the impact of writing feedback perception on self-regulated writing ability: the role of writing self-efficacy and self-regulated learning strategies. Behavioral Sciences, Vol. 15(2), 2025.

課題

受講者が多いとすべての受講者に個人的なレビューをすることが難しい（教師の負担が大きい）

解決策

peer review (受講者同士の相互評価)

2.2 peer review とは？

- 同じ立場の者同士でレビューし合うこと
- メリット
 - ▶ 批判的思考や議論のスキルを養える[3]
 - ▶ 学生同士の協力を促す[4]
 - ▶ ライティング能力を向上させる[5]

[3]Omid Noroozi, Harm Biemans and Martin Mulder. Relations between scripted online peer feedback processes and quality of written argumentative essay. The Internet and Higher Education, Vol. 31, pp. 20–31, 2016.

[4]Huda Suleiman Al Qunayeer. Supporting postgraduates in research proposals through peer feedback in a malaysian university. Journal of Further and Higher Education, Vol. 44, No. 7, pp. 956–970, 2020.

[5]O. Banihashem S.K. Biemans H.J.A. et al Noroozi. Design, implementation, and evaluation of an online supported peer feedback module to enhance students' argumentative essay quality. Educ Inf Technol, Vol. 28, pp. 12757–12784, 2023.

課題

- レビューの質が保証されない
- レビューの質が低いとライティング能力向上に効果が無い[6]

解決策？

peer review をさらにレビューして品質を向上させる

- 課題が堂々巡りになる
 - AI 導入で解決可能？

[6]Zheng Zong, Christian D. Schunn and Yanqing Wang. Do experiences of interactional inequality predict lower depth of future student participation in peer review?. Computers in Human Behavior, Vol. 127, p. 107056, 2022.

3. 研究目的

受講者が書いたレビューに対して AI から
フィードバックを与えることでレビューの質および
ライティング能力が向上するか検証する

4. 実験手法

- 実験参加者
 - ▶ 中国人の学部生 124 名（うち男性 17 名、女性 107 名）
 - ▶ CEFR B2 レベルの英語能力
 - ▶ peer review の頻度は平均して半期に一度から月に一度の間（21.77% が経験なし）
- オンラインフィードバックプラットフォーム EvaluMate を作成
 - ▶ EvaluMate 上で授業外の課題として peer review を行う

4.2 レビュー方法

4. 実験手法

The screenshot displays the EvalMate interface. On the left, under 'Evaluate this essay', there are three rows of criteria: Content, Organization, and Language. Each row has a grid of five buttons numbered 1 to 5. Below these is a 'Comment:' section containing a text box with feedback: 'Your essay is well-developed. Consider mentioning instances where machines have made mistakes or failed due to their lack of adaptability or creativity. This would further strengthen your argument about the limitations of machines. To further enhance the structure, consider using topic sentences at the beginning of each paragraph to clearly outline the main point you will be discussing. The language used in the essay is clear and precise. You have effectively made your points while avoiding repetition.' Below the text box is a '77 words' indicator and a 'Submit' button. On the right, there is a chat interface with a character named 'Eva'. The first message from Eva says: 'Let me know if you want me to help check your comments.' Below this is a green 'Check my comments' button. The second message from Eva says: 'In terms of the five aspects: 1. Affect: The feedback provided in this comment is positive and encouraging. It acknowledges the essay as well-developed and compliments the clarity and precision of the language used. 2. Description: The feedback includes a specific suggestion to mention instances where machines have made mistakes or failed due to their lack of adaptability or creativity. It also suggests using topic sentences at the beginning of each paragraph to enhance the structure. 3. Identification: The feedback explicitly identifies the lack of mentioning instances'. Below this is a blue 'Check my comments (2 times left)' button.

Figure 2: EvalMate 操作画面 出典: Kai Guo(2024)[1] より引用

- 内容、構成、言語の 3 項目をそれぞれ 5 段階で評価
- 5 つの評価基準（後述）に従いレビューコメントを記述
 - ▶ このタスクを AI がサポート

[1]Kai Guo, Mengru Pan, Yuanke Li and Chun Lai. Effects of an ai-supported approach to peer feedback on university efl students' feedback quality and writing ability. The Internet and Higher Education, Vol. 63, p. 100962, 2024.

4.3 peer review の評価基準

以下の 5 つの項目に従って peer review を評価する[7]

1. 感情(Affect)
 - ・ 励ましや賞賛が含まれているか
2. 説明(description)
 - ・ 内容の説明や、取られた行動などの要約が大きく含まれているか
3. 特定可能(Identification)
 - ・ 問題を明確かつ局所的に指摘しているか
4. 正当性(Justification)
 - ・ 特定の問題に対して、詳細な説明や正当な根拠が含まれているか
5. 建設性(Constructiveness)
 - ・ さらなる改善のための提言や行動計画が含まれているか

[7]Nafiseh Taghizadeh Kerman, Omid Noroozi, Seyyed Kazem Banihashem, Morteza Karami and Harm J.A. Biemans. Online peer feedback patterns of success and failure in argumentative essay writing. Interactive Learning Environments, Vol. 32, No. 2, pp. 614–626, 2024.

- プレテスト（week 1）

1. 以下の題材で小論文を書く（40 分）

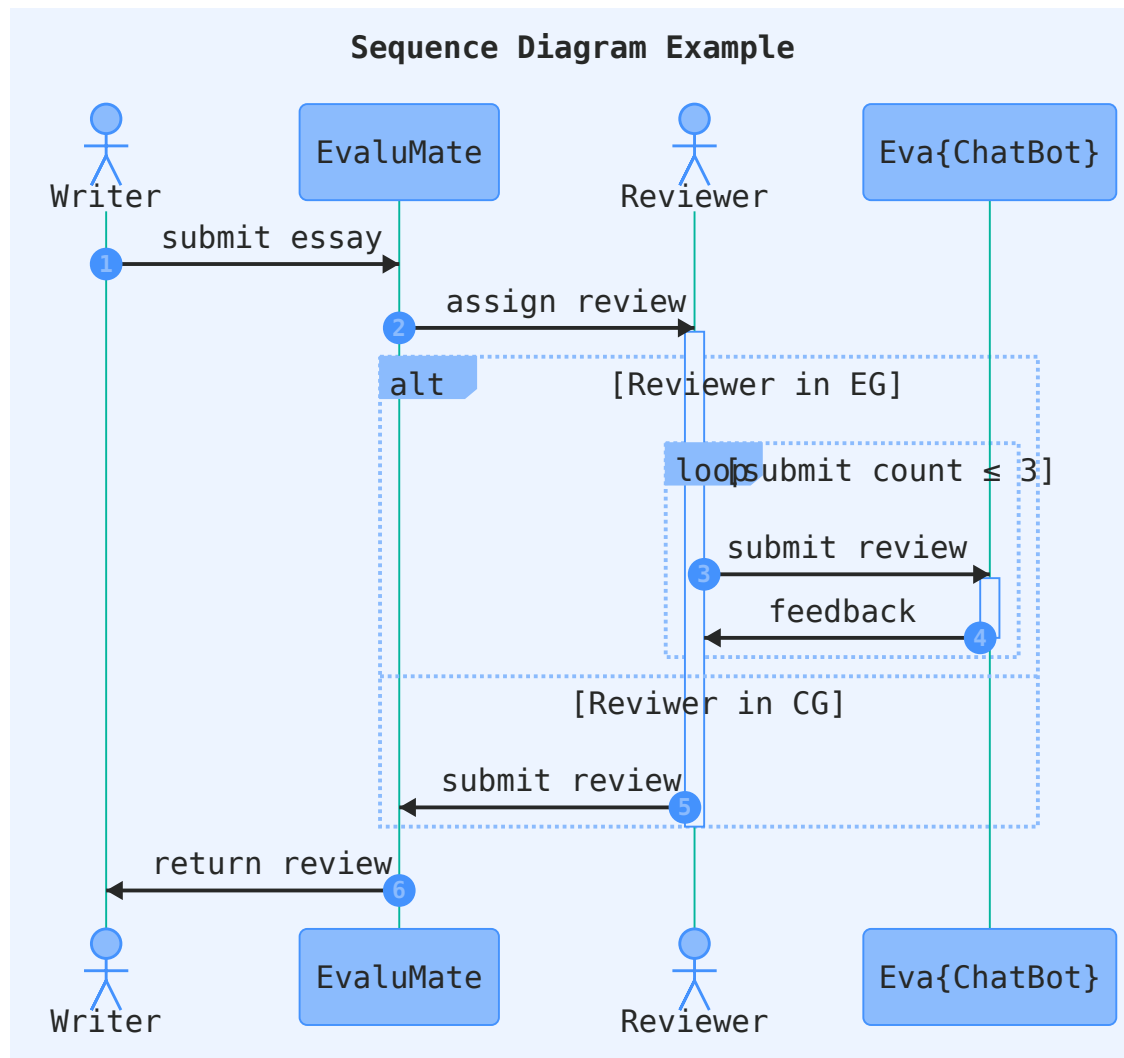
「政府はすべての学生に大学教育を無償で提供すべきである」この声明にどの程度賛成か反対かを論じ、あなたがとる立場の理由を述べなさい。あなたの立場を展開し支持する際には、関連する証拠を含めること。300-500 ワードで書きなさい。

2. ひとりの学生の小論文に対してレビューを書く（20 分）

3. レビューのトレーニング & EvaluMate のレクチャー

4.4 実験の流れ

4. 実験手法



実験 (week 2-4)

- 授業中に小論文(essay) を書いて提出
- 授業後に 3 つの小論文の review が課題として割り当てられる
- 評価基準は全体に公開されている
 - ▶ 実験群は ChatBot からのフィードバックを受けられる
 - ▶ ChatBot からのフィードバックは 1 つのレビューについて 3 回まで

4.4 実験の流れ

- ポストテスト（week 5）

1. 以下の題材で小論文を書く（40 分）

大学生は、専攻分野を選択する際、その分野の仕事があるかどうかを基準にすべきである」。この声明にどの程度賛成か反対かを論じ、あなたがとる立場の理由を述べなさい。あなたの立場を展開し支持する際には、関連する証拠を含めること。300-500 ワードで書きなさい。

2. ポストテスト時と同じ学生の小論文に対してレビューを書く（20 分）

4.5 分析方法と使用データ

- 使用データ
 - ▶ プレテスト、およびプレテストで書かれた小論文 248 本
 - ▶ プレテスト、およびプレテストで書かれたレビュー 248 本
- レビューの分析
 - ▶ 第一著者と第二著者が peer review の評価基準に従って、5 つの項目それぞれについて 2 点満点、合計 10 点満点で評価
- 小論文（ライティング能力）の分析
 - ▶ 第一著者と第二著者が内容、構成、言語の項目それぞれについて 5 点満点、合計 15 点満点で評価
- 統計分析は Quade 検定を用いた

5. 結果

5.1 レビューの質への効果

感情以外の項目で実験群が有意に質が高かった。

Table 1. Quade 検定の結果(peer review の質)出典: Kai Guo(2024)[1] より引用

Feedback quality dimension	Group	Pretest M (SD)	Posttest M (SD)	F
Affective	EG	1.91(0.29)	1.98(0.13)	0.415
	CG	1.90(0.30)	1.97(0.18)	
Description	EG	0.02(0.13)	1.20(0.82)	39.129*
	CG	0.07(0.31)	0.40(0.69)	
Identification	EG	0.95(0.58)	1.52(0.59)	31.696*
	CG	0.90(0.44)	0.85(0.71)	
Justification	EG	0.11(0.31)	1.19(0.66)	41.331*
	CG	0.15(0.36)	0.45(0.67)	
Constructive	EG	0.14(0.35)	1.45(0.71)	47.529*
	CG	0.15(0.36)	0.60(0.67)	
Totalscore	EG	3.13(0.79)	7.34(1.84)	111.508*
	CG	3.17(0.67)	4.27(1.38)	

5.2 ライティング能力への効果

内容と言語の項目で実験群が有意にスコアが高かった。

Table 2. Quade 検定の結果（ライティングスコア）出典: Kai Guo(2024)[1] より引用

Essay quality dimension	Group	Pretest M (SD)	Posttest M (SD)	F
Content	EG	2.20(0.54)	3.19(0.66)	13.372*
	CG	2.18(0.39)	2.75(0.65)	
Organisation	EG	3.19(0.53)	3.88(0.63)	0.141
	CG	2.98(0.47)	3.78(0.69)	
Language	EG	2.89(0.62)	3.97(0.53)	4.131*
	CG	2.83(0.64)	3.77(0.53)	
Totalscore	EG	8.28(1.37)	11.03(1.40)	5.124*
	CG	8.00(1.10)	10.30(1.42)	

6. 私見

- 構成の観点で実験前後で有意差がみられなかったのはレビューに構成に関する要素があまり含まれていなかったからではないか
- ただレビューを受けるだけの場合と、レビューもする場合でどの程度レビューの質とライティング能力に差が出るのか
- 実験では AI からのフィードバックを受けられる回数が 3 回だったが、ここの上限回数を増やすとより能力は向上するのか
- 小論文を読んでからレビューを返すまでの時間は実験後に早くなったのか

- AI と強調して自己調整能力を高めるシステムがあれば任意の能力を高速に高められるのでは？

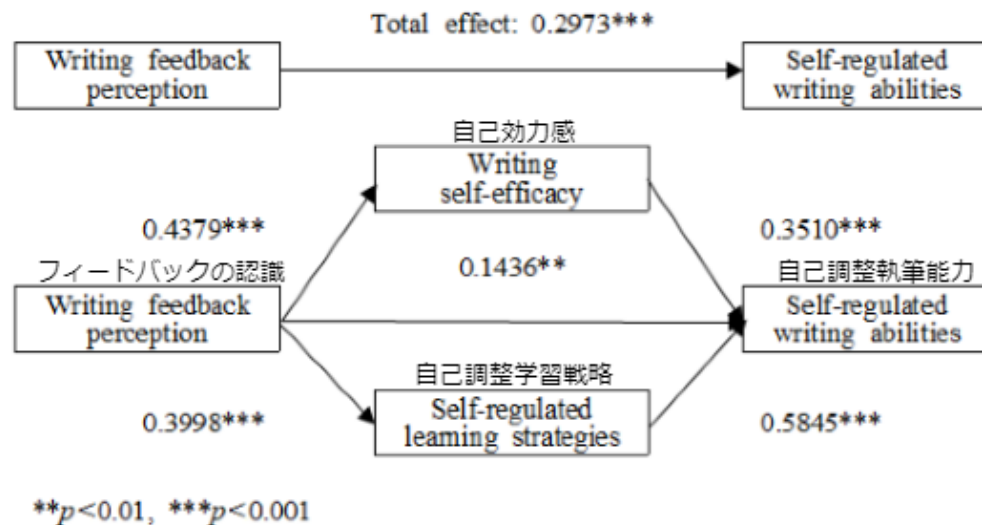


Figure 3: フィードバックが執筆能力に影響を与える経路 出典: Hwang S. (2025)[2] より引用改変

7. 参考文献

- [1] Kai Guo, Mengru Pan, Yuanke Li and Chun Lai. Effects of an ai-supported approach to peer feedback on university efl students' feedback quality and writing ability. *The Internet and Higher Education*, Vol. 63, p. 100962, 2024.
- [2] Hwang S. Unpacking the impact of writing feedback perception on self-regulated writing ability: the role of writing self-efficacy and self-regulated learning strategies. *Behavioral Sciences*, Vol. 15(2), 2025.
- [3] Omid Noroozi, Harm Biemans and Martin Mulder. Relations between scripted online peer feedback processes and quality of written argumentative essay. *The Internet and Higher Education*, Vol. 31, pp. 20–31, 2016.
- [4] Huda Suleiman Al Qunayeer. Supporting postgraduates in research proposals through peer feedback in a malaysian university. *Journal of Further and Higher Education*, Vol. 44, No. 7, pp. 956–970, 2020.
- [5] O. Banihashem S.K. Biemans H.J.A. et al Noroozi. Design, implementation, and evaluation of an online supported peer feedback module to enhance students' argumentative essay quality. *Educ Inf Technol*, Vol. 28, pp. 12757–12784, 2023.
- [6] Zheng Zong, Christian D. Schunn and Yanqing Wang. Do experiences of interactional inequality predict lower depth of future student participation in peer review?. *Computers in Human Behavior*, Vol. 127, p. 107056, 2022.
- [7] Nafiseh Taghizadeh Kerman, Omid Noroozi, Seyyed Kazem Banihashem, Morteza Karami and Harm J.A. Biemans. Online peer feedback patterns of success and failure in argumentative essay writing. *Interactive Learning Environments*, Vol. 32, No. 2, pp. 614–626, 2024.