

授業後の要約ノート作成支援アシスタントの設計と評価

香取 浩紀[†] 楊 期蘭[‡] 矢谷 浩司[§]
東京大学[†] 東京大学[‡] 東京大学[§]

1 はじめに

要約ノート学習とは、授業後に授業の内容を振り返りながらノートにまとめ直す学習である。書くことを通じて理解や記憶が深まる一方、多くの場合一人で作成することが多く、モチベーションの低下、ノートの質の低下、ノートを書くことに対する心理的コストの高さが問題として指摘されている。ノートの執筆を支援するこれまでの研究では、学習者に対する介入的重要性が指摘され、共同執筆を促進するシステムや構造化されたノートを書くことを支援するシステムが提案されてきた。しかし、他者の労力が必要であり、自習的側面の強い要約ノート作成の支援という点では、これらのシステムは汎用性を欠いている。

そこで、本研究では、生成 AI を利用したライティングアシスタントを搭載したノートシステムを実装し、アシスタントによる介入を通じた要約ノート作成の支援が効果的かを調査するユーザスタディを実施した。ライティングアシスタントは、要約ノートの内容を読み、学習者に対して質問または説明を生成する。ユーザスタディは、アシスタントが説明を生成する群（説明群）、質問を生成する群（質問群）、アシスタントなし（対照群）の 3 群の被験者間実験で、参加者は実際に録画講義を視聴して要約ノートを書き、テストを受験した。本研究で明らかにしたいリサーチクエスチョンは、以下の 2 つである。

RQ1 アシスタントの有無や介入方法の違いで、要約ノートの質や量、学習成績はどのように変化するか。

RQ2 アシスタントの有無や介入方法の違いで、要約ノートを書く過程や学習者の心理がどのように変化するか。

2 ユーザスタディ

2.1 システムの概要

参加者は Google Document 上にノートを書き、ライティングアシスタントが生成した説明または質問が赤字・太字・斜体で該当箇所に直接ノートに提示される。アシスタントは、GPT4-o をモデルとして指定し、OpenAI API の Assistants API を用いて実装した。プロンプトでは、参加者が書いている要約ノート、講義の文字起こし、既に生成された説明または質問のリストを与え、Bloom's

Taxonomy [2] の記憶・理解・応用・分析の 4 段階に基づき、講義で言及されていてかつノートに不足している事柄に関連する説明または質問を生成するように指示した。説明または質問は、ユーザが生成ボタンを押すことで生成されることに加えて、5 分に 1 個のペースで強制的かつ定期的に生成された。

2.2 ユーザスタディの流れ

実験参加者は、本学の倫理審査申請の承認後、CrowdWorks を通じて募集した。各群の人数は 10 人であり、所属する群はランダムに割り振った。参加者は、全員が成人、日本語に堪能で、講義テーマの知識がない者であり、男女各々 15 人、年齢の平均値は 36.4 歳で標準偏差は 9.80 歳であった。また、1 人当たり 2601 円の謝礼を支払った。

ユーザスタディは、著者と参加者の 1 対 1 の Zoom 上で、第 1 回と第 2 回の 2 回に分けて実施した。第 1 回では、手順の説明後、約 28 分間の録画講義動画^{*1}の視聴とアシスタントを使用しないノート取りを行い、30 分間で要約ノート（質問群と説明群はアシスタントを使用）を作成し、約 25 分間の 1 回目のインタビューを実施した。第 2 回は、第 1 回のちょうど 7 日後に、5 分間で要約ノートの復習、15 分間の講義動画を基にした多肢選択問題のテスト、約 5 分間の 2 回目のインタビューを行った。

ユーザスタディの真の目的は秘匿した上で行い、第 2 回の最後にデブリーフィングを実施した。また、生成された説明や質問を利用するか否かは参加者の任意であった。2 回のインタビューでは、オープンエンドな質問を中心とした半構造化インタビューを行い、学習や要約ノートにアシスタントが与えた影響を問う質問や参加者が取った行動を発言してもらう質問などを用意した。

3 結果

3.1 定量指標の結果 (RQ1)

ノートの量については、要約ノート作成完了後のノートの文字数では、One-way ANOVA によって、3 群の間に有意差は確認されなかったが ($F(2, 28) = 0.662, p = 0.524$)、要約ノート作成時間前後（第 1 回の 30 分間）での文字数の変化量では有意差が検出され ($F(2, 28) = 4.824, p = 0.016$)、事後検定により説明群と対照群の間のみ有意差が確認された ($p = 0.011$)。ノートの質については、トピックの網羅性と各トピックの深さによるマトリックス評価である Rubric [3] を採用した（各トピックごとに 0-3 点の点数を付け、総和を求める）。著者と著者が所属する研究

Designing and Evaluating Writing Assistants for Supporting Post-Class Summary Note Creation

[†] Hiroki Katori, The University of Tokyo

[‡] Chi-Lan Yang, The University of Tokyo

[§] Koji Yatani, The University of Tokyo

^{*1} 社会心理学入門_真島理恵_高校生のための心理学講座(日本心理学会)_1. <https://www.youtube.com/watch?v=WOKzIr4YVP4>

室の修士課程1年の学生3名の計4名で評価を実施した。量と同様に、完了後のスコア($F(2,28) = 2.27, p = 0.123$)とスコアの変化量($F(2,28) = 0.268, p = 0.767$)の両方について、One-way ANOVAからは有意差は検出されなかった。学習成績については、テストの点数に対して Kruskal-Wallis検定を用いたが、有意差は検出されなかつた($\chi^2(2) = 1.02, p = 0.600$)。

3.2 定性指標の結果 (RQ2)

3群に共通して観察された要約ノート作成中の行動としては、大半の参加者が録画講義を再度視聴しスライドや講師の発言を書き写しながら、ノートや理解の欠落点を補足していた。また、自身の記憶やノートに完全に頼った者やインターネットで検索した者が一部いた。説明群の参加者は生成された説明に対する行動として、生成された場所にそのまま放置、説明を修正して適切な場所に配置、説明を読まずに削除の3パターンに大別することができた。質問群の参加者は、生成された質問に対してできるだけ答えようとした者とほとんどの質問を削除した者に分かれた。

説明群においては、アシスタントを用いた要約ノート学習の肯定的な意見として、詳細で正確な説明そのものが役立つといったことや、見落としに気付くことができる、説明を通じて自分の理解の正しさを確認できる、復習が容易になりテストで得点できたといった学習効果に関するものが挙げられた。他にも、安心感があるパートナーのように感じたという心理的影響や、生成された説明を削る方が一から書くより速くノートを書けるといった利便性に関する言及も見られた。一方で、否定的な意見としては、既にノートに書かれていたり既に生成されたことがある説明を再度生成していくのが邪魔であり、ノートを書く行為に集中するのが妨げられるという意見が目立った。また、説明が正確なのか不安に感じたり長い文章形式の説明は理解しにくいといった内容やフォーマットに対する意見、説明が教科書的でそれをコピーアンドペーストするだけでは頭に入りにくいといった学習感覚に対する意見、生成される説明に意識が向きすぎて、説明が生成されなかつ箇所の理解や記憶、ノートの記述が手薄になったという意見が得られた。

質問群においては、肯定的な意見は学習効果と心理的影響の2つに大別できた。学習効果については、説明群と同様に見落としに気付けるという意見に加え、質問に答えようと調べたり思考する過程で理解や記憶が深まる、まとめるスキルやノートの質が向上したという意見が見られた。心理的影響については、アシスタントがモチベーションを上げる役に立つ存在に感じたという発言があった。他方、否定的な意見として、説明群と同様に、既に生成されたものや答えがノートに書かれている質問が出てくるという質の低さに言及する意見や質問が次々に生成されることで集中が妨げられる、質問が生成されなかつ箇所の理解が曖昧になったといったことが得られた。加えて、要約ノートを書く時間制限がある中で、さらに質問が生成されることにより、急がされているようで焦ったという心理的

な負荷を感じた参加者も観察された。

4 考察

まとめノート作成時にライティングアシスタントを使用すべきか否か、使うとして説明形式が良いのか質問形式が良いのかについて、定量指標では多くの指標でグループ間の有意差が見られなかったことから、断定的な見解を述べることは難しい。一方で、インタビュースクリプトの定性的な分析から、アシスタントや説明・質問形式の利点や欠点、学習者の行動や知覚の変化や違いが観察された。

これらを踏まえて、まとめノート作成のライティングアシスタントの設計においては次に示す内容とインターフェースにおける指針が重要だと考える。内容の観点では、使用者にとって新規の情報を提供し続け、使用者が触れる情報とノートの網羅性を高める説明または質問を生成することが学習効率とモチベーションの両面で望ましい。また、箇条書きなど使用者にとって読みやすい形式で提示することで認知的なコストの低減に繋がると考えている。インターフェースの観点では、説明または質問はノートの紙面の横などに提示したり、生成時に通知するなどして、使いやすさを意識した設計が、参加者の気を散らさずに集中力を維持するために重要である。

説明群と質問群において、AIへの依存性を高める参加者が観察された。これは、説明または質問の生成箇所以外の学習が疎かになったという意見、ノートそのものをAIに生成して欲しいという意見、AIが生成した説明をただコピーアンドペーストしたといった行動から窺うことができる。AIの説明形式は質問形式より使用者が受動的になる傾向があり[1]、まとめノート作成でも定性的には似た傾向が見出された。AIという利便性を前に人は安きに流れがちだが、一定の認知負荷が必要な学習、特にまとめノート学習という文脈においては、AIを用いることによる利便性や依存性とAIという他者視点による学習効果の追求やモチベーションの向上のバランスをとることが肝要である。上述した指針を踏まえたアシスタントの設計、複数の教科や長期的な学習効果、認知的コストの定量評価も含んだユーザスタディが次の研究の方向性だと考える。

謝辞

ユーザスタディの参加者の方々には貴重な意見をいただき、研究室の皆様からは研究に関連する助言を多数いただきました。この場を借りて、感謝申し上げます。

参考文献

- [1] Danry, V. et al.: *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, pp. 1–13 (2023).
- [2] Jones, K. O. et al.: *2009 39th IEEE Frontiers in Education Conference*, pp. 1–6 (2009).
- [3] Peverly, S. T. et al.: *Journal of Educational Psychology*, Vol. 99, No. 1, p. 167 (2007).