

査読論文

生成 AI を活用したレポート自動採点のプロンプト設計と評価

薦 田 勇 智・宮 崎 剛 司

要約

本研究は、大学教育における生成 AI 活用の一例として、レポート自動採点のプロンプト設計と評価を行ったものである。ゼミ所属学生7名のレポートを対象に、人間の採点と Google NotebookLM を用いた AI 採点を比較した。評価は「文と語句」「段落」「全体構成」「内容の示し方」「引用と参考文献」の5観点で実施した。その結果、AI は平均点が高く、特に低評価を避ける傾向が確認された。これは AI が表面的な整合性を重視し、論理的欠陥や内容の深度を十分に捉えきれないためと考えられる。一方で、AI は一貫した基準で採点できる点から教育評価の補助的手段として有用性が示唆され、今後はプロンプト最適化や評価基準の多層化が求められる。

キーワード：生成 AI, プロンプト設計, 自動採点, 自然言語処理

1 はじめに

近年、生成 AI (Generative AI) の進展は目覚ましく、社会における浸透度はもはや「導入するか否か」を議論する段階を超えている。現在求められているのは、生成 AI をいかに制御し、教育・研究・実務の場で適切に活用するかという実践的な課題である (Brynjolfsson & McAfee, 2017)。特に大学教育の現場において、教員は学生に対して利用の意義とリスクを的確に伝えるとともに、自らも効果的な活用のモデルを提示する責務を負っている。

本研究では、生成 AI を活用するうえで不可欠な三つの基盤的技術を提示する。第一に、プロンプト構築である。生成 AI は入力される指示文 (プロンプト) の精度や工夫に強く依存し、同一の問いに対しても指示の仕方によって出力が大きく変容する (Reynolds & McDonell, 2021)。したがって、試行錯誤を通じて望ましい出力を導く能力は必須である。第二に、批判的思考である。生

成 AI は時に「ハルシネーション」と呼ばれる誤情報を提示するため、利用者は出力内容が無批判に受け入れるのではなく、必ず検証を行う必要がある (Ji et al., 2023)。近年、SNS 上で「ファクトチェック」の実践が一般化しつつあることは、その意識の広がりを示している。例えば、X (旧 Twitter) では、搭載された AI 機能 (grok) を用いてニュースや投稿の真偽を確認する取り組みが増加しており、情報の信頼性を精査する文化が徐々に根つきつつある。第三に、活用力である。生成 AI の真価は、単なる知的道具としての利用にとどまらず、利用者の目的に応じて柔軟に使い分けられる点にある。日常的な対話支援から専門的な業務補助に至るまで、その応用範囲は極めて広い。しばしば懸念される「AI による職業代替」も、歴史的には技術革新に伴う社会構造の変化の一形態に過ぎず、人間は思考・判断といった AI の限界を超える領域に適応することで、新たな役割を担ってきた (Autor, 2015)。

要するに、生成 AI を使いこなす上で必要とさ

れるのは、「伝える技術（プロンプト構築）」「見極める技術（批判的思考）」「活かす技術（活用力）」という三つの技能である。本研究はこれらの技術的基盤をふまえ、大学教育の文脈に即した生成AIの利用法を探求する。具体的には、大学生が作成したレポート文を対象とし、自動採点を可能にするプロンプトを構築するとともに、その有効性を検証することを目的とする。

2 先行研究

生成AIを活用したレポートの自動採点に関する論文では、主に以下の4本が見つかった。

No.1 岩澤ほか (2024)「生成AIを用いた自由記述の自動採点支援システムの試作と構築¹⁾」では、教師の校務DX推進のため、生成AIを活用した自由記述の自動採点支援システムを開発し実証実験を行っていた。このシステムは、教師が設定する重点ワードのカウント・ハイライト・スコア化、およびAIによる生徒回答の要約・フィードバック自動生成が特徴であり、教師の採点時間削減と的確なフィードバック提供を目指し、重点ワード活用による採点の客観性と正確性の向上を期待されていた。今後の課題では、生成AIの最新状況に合わせてシステムの連携部分を随時見直し、改善していく予定とあった。

No.2 岩崎ほか (2024)「文章校正における言語生成AIとループリックによる個別フィードバックの活用²⁾」では、大学の初年次生がレポートを校正する際に、言語生成AIとループリックを用いた個別フィードバックの活用効果と課題を検証していた。その結果、65.4%の学生が言語生成AIによるフィードバックを肯定的に評価していた。学生はフィードバックを通して、自分で気づかなかった表記表現や文章構成の問題点を把握し、改善に役立てた例が見られていた。しかし、AIのフィードバックに誤りがある点が課題として指摘され、継続的な調査の必要性が示唆されていた。

No.3 溝上ほか (2025)「ChatGPTによるループ

リック生成及びレポート自動採点の検討—アントレプレナーシップ教育におけるループリックの活用—³⁾」では、ChatGPTによるループリック生成とその自動採点の利用可能性を信頼性と妥当性の観点から検討していた。ChatGPTで生成した2種のループリックを用い、大学生のレポートを教師とChatGPTで採点を行う。修正後ループリックを用いたChatGPTの採点では高い一貫性(ICC 0.82)が示され、特に上位レポートで高い信頼性が得られていた。教師との比較では、上位グループの採点結果は同等であったが、中位・下位グループではChatGPTの得点が高くなる傾向が見られていた。結果として、ChatGPTは十分利用可能なループリックを生成できるが、自動採点への利用は現時点では難しいと判断されていた。

No.4 矢野 (2023)「OpenAI Chat APIを用いた自動採点付き学習アプリの開発と授業実践⁴⁾」では、OpenAI Chat APIを用いた短文回答形式の自動採点システムを開発し、大学のプログラミング授業で導入した。このシステムは模範解答との類似性に基づきAIが答案を評価するもので、一定の精度で採点が可能であることが示されていた。しかし、一部で誤判定(例:空白回答を正解とする、問題文のキーワードに引きずられる)も見られたため、今後はプロンプトの改善やキーワードマッチングなどの非AI分析手法との組み合わせにより、システムの効率と精度向上を目指していくとあった。

3 研究方法

筆者のゼミに所属する7名の学生(A~G, 経営学科1年生)を対象に、同一のレポート課題に対する評価を、人間による採点とAIによる採点で比較することを目的とした。レポートの条件として、次のような内容を指定した。

「eスポーツがオリンピック競技として正式に行われることになりました。

このことについて以下の内容でレポートを書いてください。

- ・ニュースでどのように扱われているのか
- ・コメントや SNS 等で見られる一般の意見はどのようなものが多いのか
- ・それらに対する自分の考え

また、条件は以下の通りです。

- ・1500 文字以上書く
- ・表紙不要、ヘッダーに学籍番号と氏名、フッターにページ番号を振る（方法が分からない場合は調べる）
- ・ニュースを第 1 章、一般の意見を第 2 章、自分の意見を第 3 章とすること
- ・最後に参考文献を書くこと（ネットの記事なので確認した年月日を書く）」

課題は 1 日遅れが一人いたものの全員提出しており、全て条件を満たしていた。また、レポートの文字数は最大で約 3500 文字、最小で約 1500 文字となっていた。そのため、これらについての評価で減点をする事は無かった。

まず、筆者自身が各レポートを従来の基準に基づき採点したうえで、AI 採点を実施した。AI 採点には、Google が提供する NotebookLM を用い、PDF 形式のレポートを読み込み、指定したプロンプトに従って評価を行わせた。

採点基準は、先行研究を参照し、以下の 5 観点を設定した。各観点は 20 点満点とし、総合得点は 100 点とした。

文と語句：文章の明確性、一貫性、正確性の総合評価

段落：段落構成の論理性と情報密度の適切性

全体構成：文章全体の構造的整合性と論理展開の明瞭さ

内容の示し方：主張の明確さと論証の妥当性

引用と参考文献：引用の適切性および学術的信頼性

AI 採点に用いたプロンプトは、これらの観点を明示し、各項目について 20 点満点で評価する

よう指示した。これにより、人間と AI の評価結果を同一基準で比較可能な形に整えた。実際のプロンプトは以下のとおりである。

「このループリックを使って任意の大学レポートを採点してください。各観点ごとに、文の明確さ、段落構成、全体構成、内容の充実度、引用の適切さを評価し、良い点と改善点を具体的に記述してください。最終的に総合評価をつけ、学生にわかりやすいフィードバックを作成してください。さらに、複数のレポート形式（論説文、研究報告、ケーススタディ等）に対応できるようにしてください。必要に応じて評価基準を微調整してください。

評価手順：

文と語句：一文一義で明確か、専門用語や定義が適切に説明されているか、誤解を招く表現がないかを確認します。また、文体の統一性、表現の正確さ、文法や句読点の使用も評価対象に含めます。段落：各段落に中心文があるか、段落内の文が中心文を支えているか、段落同士の論理的なつながりを評価します。段落の長さや情報密度もチェックし、冗長な文章や飛躍のある展開を指摘します。全体構成：序論・本論・結論が明確に分かれているか、序論で提示した目的や問いが本論で体系的に論じられているか、結論で適切にまとめられているかを確認します。必要に応じて見出しや章立ての適切性も評価します。

内容の示し方：主張が明確で、証拠やデータを用いた論証がされているか、課題に沿った内容か、独自の視点や分析が含まれているかを評価します。また、情報の正確性、論理の一貫性、結論の妥当性も確認します。

引用と参考文献：引用の正確性、文中での引用の文脈適合性、出典提示、参考文献リストの形式が適切かを評価します。引用の信頼性、最新性、関連性も確認し、適切な学術的基準に従っているかをチェックします。

出力フォーマット：

観点ごとの評価：良い点／改善点／具体例／20
点満点中何点か（5点刻み）

総合評価：（100点満点）とその理由

フィードバック文：学生向けに、具体的で建設的
かつ実践的な改善提案を含むコメント

適用レポート形式：評価対象のレポートが論説文、
研究報告、ケーススタディなどどの形式か明示し、
それに応じて採点基準を調整する

追加指示：

各観点ごとに最低1つの具体例を示すこと

良い点だけでなく、改善点を明確に指摘すること

総合評価は観点ごとの評価を総合して導くこと

学生が次回改善できるように、実践的で具体的な
アドバイスを含めること

採点結果は簡潔で読みやすく、教育的価値を重視
すること

必要に応じてレポート形式やテーマに合わせた補
足的な評価基準を追加すること」

4 結果・考察

上記の観点から、筆者が行った場合の各学生の
採点結果は以下の通りになった。

表1 筆者が行った場合の各学生の採点結果

	内容の示し方	段落	引用と参考文献	全体構成	文と語句	合計
A	15	15	20	5	10	65
B	5	10	5	5	5	30
C	20	15	10	10	15	70
D	20	15	10	10	15	70
E	10	15	15	10	10	60
F	10	10	5	5	15	45
G	15	15	20	10	15	75

表1に示す通り、最高点70点、最低点30点、
平均点は60点であった。注目すべきは、いずれ
の学生も全項目において満点を獲得するには至ら
ず、特に「全体構成」の評価項目においては、点
数を付けることが困難であった点である。この傾
向は、対象が大学1年生であるという事実と密接
に関係していると考えられる。すなわち、レポー
ト作成における構成力の未成熟さが、評価の難し

さとして顕在化したのである。

一方で、「段落」および「内容の示し方」に関し
ては、比較的高い評価が可能であった。これは、
文章の基本的な構成要素に対する理解が一定程度
形成されていることを示唆している。筆者自身、
採点基準を意図的に緩和した認識は持っていない
が、評価対象が大学1年生であるという文脈を踏
まえた結果として、相対的に寛容な評価となった
可能性は否定できない。

次に、生成AIによる採点結果は以下の通りにな
った。

表2 生成AIが行った場合の各学生の採点結果

	内容の示し方	段落	引用と参考文献	全体構成	文と語句	合計
A	15	15	5	15	15	65
B	10	10	5	10	15	50
C	15	15	5	10	15	60
D	15	15	5	15	15	65
E	15	15	5	15	10	60
F	15	15	5	15	15	65
G	18	18	10	15	15	76

最高点76点、最低点50点、平均点は63点と、
筆者による採点結果と比較して、全体的に高め
の傾向が見られた。特筆すべきは、点数の分布が極
端な偏りを示さず、均質性が高いことである。こ
れは、AIが採点において一貫した評価基準を適
用していることの表れであるが、同時に、レポー
ト間の質的差異を十分に反映しきれていない可能
性も示唆する。

これらの結果をそれぞれの学生に応じてレー
ダーチャートで纏めた。すると、図1のようにな
った。

レーダーチャートの形状を比較すると、学生間
で評価の乖離パターンに顕著な差異が認められ
た。評価の一致度に基づき、対象学生を3群に分
類すると、まず学生BおよびFにおいては、AI
評価が筆者評価を大幅に上回り、レーダーチャー
トの形状が著しく異なる結果となった。特に学生
Bは筆者評価30点に対しAI評価50点、学生F
は筆者評価45点に対しAI評価65点と、それぞ
れ20点の差が生じており、これらの学生のレポー
トは、筆者が「論理的一貫性の欠如」「主張の根拠
不足」「構成の未熟さ」といった本質的な問題点を

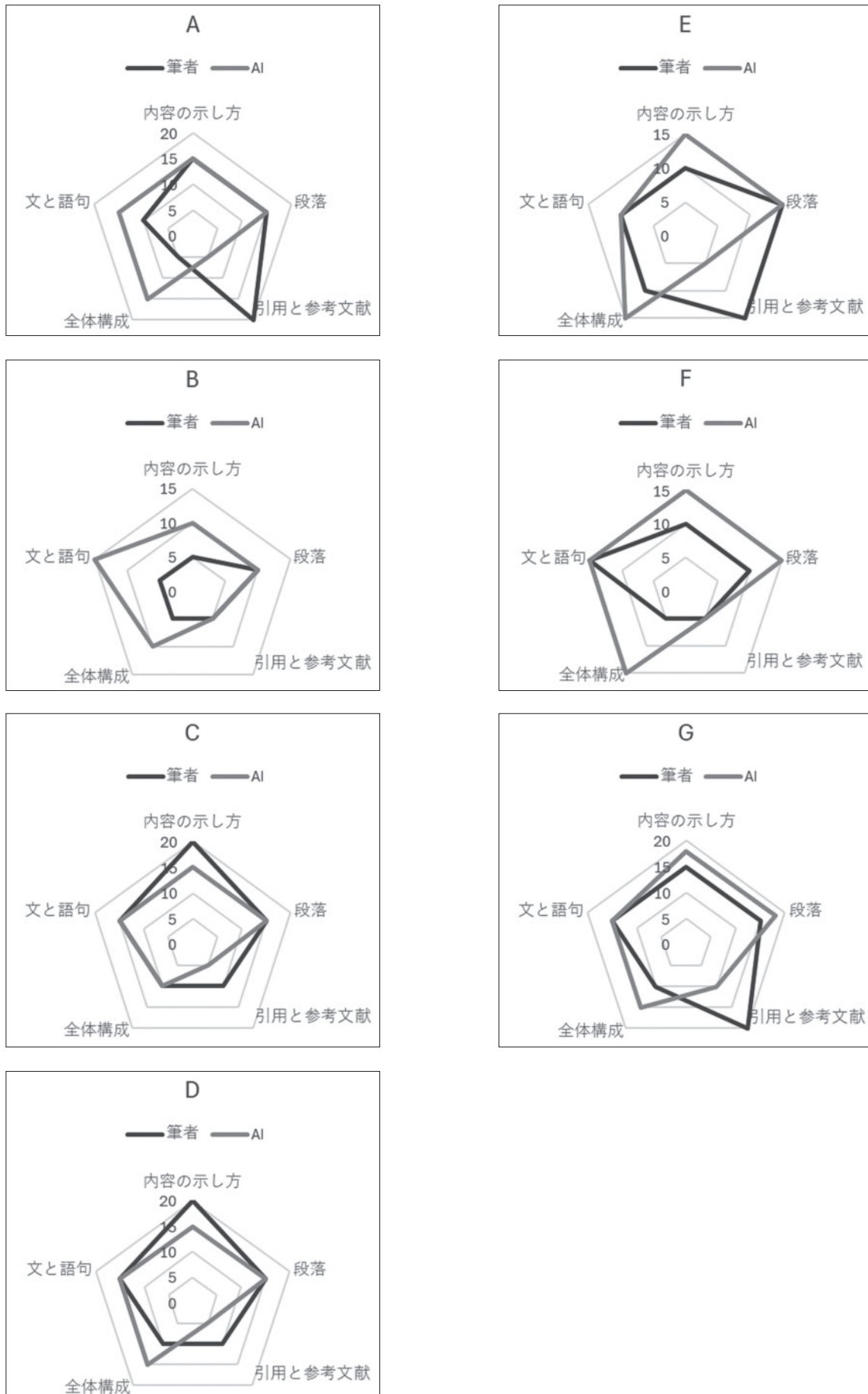


図1 筆者とAIによる観点別の採点差

指摘したのに対し、AIは表面的な文章の整合性や形式的要件の充足を評価した結果、大きな乖離が生じたと考えられる。次に、学生A、E、Gにおいては、一部の観点でAI評価が筆者評価を上回る一方、他の観点では比較的近接した評価を示すという混合的なパターンが観察され、具体的には「内容の示し方」「全体構成」といった要素ではAIが高めの評価を付与する一方で、「文と語句」の要素では両者の評価が接近する傾向が見られた。一方、学生CおよびDにおいては、筆者評価とAI評価が全観点にわたり比較的近接しており、レーダーチャートの形状も類似していた。このような評価の乖離パターンの多様性は、AIによる評価が、レポートの質的水準によって異なる妥当性を示すことを意味しており、すなわち、一定の水準を満たしたレポートに対してはAIは妥当な評価を下せるが、質的に問題のあるレポートに対しては、その問題点を十分に捉えきれず、過大評価する傾向があるということである。

特に学生Bおよび学生Fに関しては、筆者による評価とAIによる評価の間に最大20点の乖離が生じた。筆者は両者に対して、それぞれ30点および45点という低評価を与えているが、AIはそれに対して相対的に高い点数を付与している。このような現象は、先行研究においても指摘されている通り、生成AIが極端に低い評価を避ける傾向を持つことと一致する(溝上 *et al.*, 2025)。この傾向は、AIの設計思想に内在する「中庸性の志向」に起因する可能性がある。

学生BのAIからのフィードバック⁵の一部を抜粋すると「テーマに関する情報の収集 (IOCの取り組み、賛否両論の整理) は行えているものの、学術レポートとして最低限守るべき引用ルールが著しく欠如している (5点)。また、結論の章がないという構成上の大きな問題があり、論説文としての役割を果たせていない。」また、学生FのAIからのフィードバックの一部を抜粋すると「学術レポートとして必須である「引用と参考文献」のルール (本文中での出典明記) が全く守られていないという形式的欠陥 (5点) が、信頼性を著し

く損なっています。また、自身の主張が、本論で提示した倫理的な課題への反論を欠いているため、論理的な一貫性にも課題が残ります。」とある。他の学生でもそうだが、必ず「引用と参考文献」のルールについて触れている。この現象の背景には、本研究で用いたレポート課題の特殊性が影響していると考えられる。今回の課題テーマは「eスポーツのオリンピック正式種目化」であり、極めて時事性が高く、かつ学術論文として十分に蓄積されていない新興的なトピックである。そのため、筆者は課題実施時に、学術論文に加えて信頼性の高いWebサイト (例: 国際オリンピック委員会の公式声明、主要メディアの報道記事、業界団体の調査報告書など) の引用を許可していた。

しかしながら、学術論文執筆の規範に照らせば、査読付き学術誌や専門書からの引用が望ましいことは言うまでもない。筆者は、この理想と現実のギャップを認識した上で、各学生のレポートにおける引用元の信頼性、引用の文脈適合性、出典表記の正確性を厳格に評価した。その結果、Webサイトのみに依拠した引用、出典が不明瞭な記述、引用と本文の論理的結びつきの弱さなどが複数のレポートで確認され、相対的に低い評価となった。

一方、AIは引用元の媒体 (学術論文かWebサイトか) よりも、形式的な整合性、すなわち引用表記の有無、参考文献リストの存在、引用番号の対応関係などを重視して評価を行った可能性が高い。このことは、AIが「形式的正確性」を優先し、「学術的信頼性」という質的側面の判断において限界を持つことを示している。

ここで一度立ち止まり、評価という行為そのものの性質について再考する必要がある。人間による評価は、文脈、経験、そして瞬間的な印象に左右される可変的な営為であるのに対し、AIによる評価は、定量的な特徴抽出とアルゴリズムの処理に基づく、ある種の「構造化された判断」である。この差異は、教育現場における評価の在り方に対して、根本的な問いを投げかけるものである。

すなわち、「評価の客観性」とは何か、「公正な採点」とはどのように定義されるべきかという問

題である。本研究は、AI と人間の評価の差異を単なる数値の比較として捉えるのではなく、その背後にある認知的枠組みの違いを浮き彫りにすることを目的としている。今後の研究においては、AI による評価が教育的介入にどのような影響を及ぼすのか、また、評価の質的側面をどのように担保するのかという課題に対して、より精緻な検討が求められるだろう。

5 今後の課題

本稿における検証を通じて明らかとなったのは、生成 AI によるレポート評価において、特に低評価の付与に関して一貫性を欠く傾向が存在するという点である。先行研究（溝上 *et al.*, 2025）でも指摘されているように、生成 AI は高評価のレポートに対しては比較的妥当な点数を提示する一方で、低評価が想定されるレポートに対しては、実際よりも高い点数を与える傾向が観察された。

この現象は、AI の評価アルゴリズムが、文章の構造的整合性や表面的な言語的流暢さを過度に重視する一方で、内容的な深度や論理的欠陥といった、より本質的な評価基準を十分に捉えきれていない可能性を示唆している。したがって、プロンプト設計において、評価の幅を意図的に拡張するような指示、すなわち「優劣の差異を明確に反映させる」ような条件設定を行うことで、より意図に近い評価結果を得ることが可能となると考えられる。

この知見は、生成 AI を教育評価の補助的手段として活用する際の設計指針に対して、重要な含意を持つ。すなわち、AI の評価能力を盲目的に信頼するのではなく、その限界を冷静に見極めた上で、評価プロセス全体を人間の判断と補完的に設計する必要がある。今後の課題としては、AI による評価の精度を高めるためのプロンプト最適化手法の体系化、および評価基準の多層化に向けた理論的枠組みの構築が求められる。

参考文献

- ・寺岡三左子（2023）. ChatGPT® は有効な教育支援ツールになり得るか？—看護学における対話型 AI をめぐる議論の動向—, 医療看護研究, 20(1), 64-70
- ・新原俊樹, 甲斐尚人, 小柏香穂理, 船越幸夫（2024）. ChatGPT を活用した研究データの作成事例, 情報知識学会誌, 34(1), 18-23
- ・小原有以, 佐藤美唯, 倉光君郎（2023）. KOGI : ChatGPT を Colab に統合したプログラミング演習支援, 情報教育シンポジウム論文集, 2023, 141-148
- ・溝上広樹, 中山泰宗（2025）. ChatGPT によるループリック生成及びレポート自動採点の検討—アントレプレナーシップ教育におけるループリックの活用—, 崇城大学紀要, 50, 75-84
- ・Mohanad Halaweh（2023）. ChatGPT in education: Strategies for responsible implementation, Contemporary Educational Technology, 15(2), ep421
- ・Houda Bouamor, Juan Pino, & Kalika Bali（2023）. Is ChatGPT a General-Purpose Natural Language Processing Task Solver?, Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing, Volume 1 of 19, 1339-1384
- ・Impact of ChatGPT on Learning Motivation

文末注

- 1) 岩澤孝徳, 久保田真一郎, 喜多敏博, マジュンダール・リトジット（2024）, p.685-686
- 2) 岩崎千晶, 松河秀哉（2024）, p.1-4
- 3) 溝上広樹, 中山泰宗（2025） p.75-84
- 4) 矢野（2023） p.25-31
- 5) 付録に各学生に対するフィードバックの全文を記載

付録

学生 A へのフィードバック
総合評価：65 点（100 点満点）

その理由：
本レポートは、e スポーツのオリンピック参加というテーマに対し、現状の背景（OEG 決定）、主要な批判、そして自身の見解という流れで議論を試みており、

構成の意図は明確です。特に、スポーツ基本法や平林氏の問いを参照し、論争の核心を捉えようとしている点は評価できます。しかし、学術レポートの根幹である「引用と参考文献」のルール（本文中での出典明記）が全く守られていない（5点）ため、形式的な信頼性が著しく低くなっています。また、自身の主張が主観的な経験論に留まり、議論で提起された著作権や普遍性といった形式的な課題に対する反論が不足しているため、内容的な深掘りが今後の課題です。

フィードバック文（学生向け）

件名：レポート「オリンピック競技となったeスポーツについて」に関する評価と具体的な改善提案

レポート作成お疲れ様でした。ご自身のバドミントンや空手の経験、そして熱心なゲーム経験に基づいてeスポーツの魅力を論じようとした点は評価できます。特に、平林氏の疑問やスポーツ基本法の定義といった客観的な情報を用いて、議論の対立軸を明確に示そうとした努力は素晴らしいです。

しかし、このレポートをさらに高いレベルに引き上げ、今後の学術的な文章作成で「文章の癖」を修正するために、以下の点を次回以降必ず実践してください。

1. 学術的引用ルールの徹底遵守（最重要課題）：

実践的な改善提案：あなたのレポートは、IOCの決定、平林氏の疑問点、スポーツの定義など、外部の情報源を多数利用しています。しかし、どの情報がどこから来たのかを本文中で全く示していません。この「癖」を修正するため、次回以降は、事実や定義、他者の意見を記述した際は、その直後に必ず（出典名、年）や（メディア名）といった形式で出典をカッコ書きで明記してください。これを行うことで、あなたのレポートの信頼性は飛躍的に向上します。

2. 言葉の繰り返しと接続詞の修正：

実践的な改善提案：文章内で「私自身」や「～である」といった表現を不必要に繰り返す癖が見られます。また、「だか」のような不適切な接続詞の使用も散見されます。今回は、一度使った言葉や主語を繰り返さずに、より自然な代名詞やより適切な接続詞（例：「しかし」「一方で」）を選び、文章の流れをスムーズにする訓練をしてください。特に「だか」は「だが」に直す必要があります。

3. 議論の論理的な深掘り：

実践的な改善提案：あなたの最終的な主張（eスポー

ツの採用は間違いではない）は、スポーツと共通する「経験」（挫折、努力、チームワーク）に強く依存しています。これは説得力がありますが、平林氏が提起した「私企業の著作物であること」「普遍性の欠如」という、オリンピックの理念と公平性に関わる形式的な問題に対し、あなたの経験論がどう反論し、解決できるのかを具体的に論じてください。「経験が共通しているからOK」ではなく、「経験が共通している上に、これらの形式的な問題は、IOCの新たな枠組み（OEG）によってこう解決できる」という論理的な裏付けを加えることが、内容充実度を高める鍵です。

学生Bへのフィードバック

総合評価：50点（100点満点）

その理由：

本レポートは、テーマに関する情報の収集（IOCの取り組み、賛否両論の整理）は行えているものの、学術レポートとして最低限守るべき引用ルールが著しく欠如している（5点）。また、結論の章がないという構成上の大きな問題があり、論説文としての役割を果たせていない。自身の主張は賛成であると明確に述べているが、その論拠が浅く、反対意見で挙げられた重要な課題（著作権、倫理）に対する論理的な反駁や解決策が提示されていないため、全体的な説得力に欠けている。

フィードバック文（学生向け）

件名：レポート「eスポーツオリンピックについて」に関する評価と改善提案

レポート作成お疲れ様でした。eスポーツのオリンピック参加という現代的なテーマを取り上げ、IOCの具体的な動きや、賛成・反対両方の主要な論点を整理しようとした努力は評価できます。

しかし、このレポートを大学レベルの学術的な文章に引き上げるためには、以下の3点について、あなたの文章作成の「癖」を具体的に修正する必要があります。

1. 引用ルールの徹底遵守（形式面の最重要課題）：

実践的な改善提案：あなたのレポートは、多数の外部情報（ニュース、他者の意見など）に基づいているにもかかわらず、本文中で「誰が」「いつ」述べた情報を指し示す文中引用（インテキスト・サイテーション）が全く行われていません。これは、学術的な厳密さの観点から非常に大きな問題です。次回以降は、外

部情報を使用する際は必ずその直後に（出典名，年）などの形式で出典を示し，レポートの説得力を保証してください。

2. 論理構成の完成（結論の作成）：

実践的な改善提案：現在のレポートには「結論」の章がありません。論説文では，導入で提示したテーマに対し，本論で展開した議論（賛否両論）を総括し，最終的に自分の主張の妥当性を再確認する「結論」の章が必須です。今回は，第3章を「結論」とし，自分の賛成意見が，反対意見（著作権，暴力性など）の懸念をどう乗り越えようのかという視点を含めて，レポートを力強く締めくくる文章を作成してください。また，レポートの途中に AI に作成させたと思われる「AI の意見」を挿入することは，レポートの著者の主張を曖昧にするため避けてください。

3. 議論の深掘り（主張の論証）：

実践的な改善提案：あなたの主張（e スポーツに賛成）は明確ですが，根拠が「知名度が上がる」「チャンスになる」といった情緒的・一般的なメリットに留まっています。反対意見（著作権，ゲームの暴力性）は，オリンピックの理念に関わる重い論点です。今回は，これらの論点に対し，「ではどう解決できるのか」という理論的・具体的な反論を展開してください。例えば，暴力的なゲームを使わない，あるいは著作権問題を解決する新たな組織構造を提言するなど，論点を深掘りする癖をつけてください。

これらの改善は，学術的な文章を書く上で非常に重要です。特に引用のルールを習得することが，あなたのレポートの信頼性を高める第一歩となります。

学生 C へのフィードバック

総合評価：60 点（100 点満点）

その理由：

本レポートは，e スポーツのオリンピック参加に対する反対意見を，伝統，身体性，著作権など複数の観点から論じようとしており，主張は一貫しています。しかし，論説文の根幹となる全体構成の弱さ（結論が先行し，最終的なまとめがない），および学術的な基礎である引用と参考文献の欠如が大きな減点要因となりました。議論のアイデアは良いものの，それをレポートとして形式化し，論理的に展開するスキル，特に客観的な情報への適切な対応に改善の余地があります。

フィードバック文（学生向け）

件名：レポート「e スポーツのオリンピック参加について」に関する評価と改善提案

今回のレポート作成お疲れ様でした。e スポーツのオリンピック参加というタイムリーなテーマに対して，伝統性，肉体性，著作権の問題など，多角的な視点から反対意見を構成しようとした点は非常に評価できます。特に，国内で反対意見が依然として多数派であることを示す世論調査の具体的なデータ（日テレ系『スッキリ』の投票結果）を盛り込んだ点は，説得力を高める良い試みです。

しかし，次の点について改善することで，レポートの質を大きく向上させることができます。

1. 構成の見直し（結論の位置）：

実践的な改善提案：現在，第3章の冒頭で「反対と考える」と結論を先に述べてしまっています。これは読者に対して強い印象を与えますが，論説文としては，序論（テーマ提示）→本論（複数の論拠と証拠）→結論（主張の再確認と総括）という構造を守るべきです。今回は，本論で論拠を十分に展開した後，結論の章で改めて自分の主張の妥当性を強調する構成に修正してください。レポートの最後が，引用に関する障害の指摘で終わるのではなく，全体の主張を締めくくる段落で終わるようにしてください。

2. 引用ルールの遵守：

実践的な改善提案：大学のレポートでは，外部から得た情報（ニュース，統計，他者の意見，専門的な知識）全てに，出典を明確に示さなければなりません。今回のレポートでは，IOC の決定や世論調査の結果，著作権に関する議論など，多くの外部情報を参照しているにもかかわらず，正式な引用形式（例：[著者名，年]）が使われておらず，参考文献リストもありません。次回以降は，引用した情報全てについて，カッコ書きで出典を示し，レポート末尾に必ず参考文献リストを作成してください。

3. 議論の深掘り（「癖」の修正）：

実践的な改善提案：賛成意見に対する反論で，「精神的スキルや戦略は多くの伝統的スポーツにも共通している」と指摘していますが，この反論はやや弱く，議論を深める機会を逃しています。今回は，相手の論点を正面から受け止め，「e スポーツの戦略性は伝統スポーツのそれと比べて具体的にどう異なるのか，そしてなぜそれがオリンピック競技として不適格なのか」という分析を加えることで，議論を深めてください。

4. 文体の改善:

実践的な改善提案:「国際オリンピック委員会 (IOC)」のように、一度説明した語句をそのまま繰り返さず、二度目以降は略称 (IOC) のみを使うなど、冗長な表現を避ける工夫をしてください。また、一つの文章に複数の情報や判断を盛り込みすぎないように、「一文一義」を意識して文を短く区切る訓練をすると、文章の明確さが飛躍的に向上します。

学生 D へのフィードバック

総合評価: 65 点 (100 点満点)

その理由:

本レポートは、e スポーツのオリンピック参加に対する賛成意見を、既存の批判を踏まえて論じようとした点は評価できます。特に、反対意見の背景に「年齢層による偏見」があるという分析は興味深いです。しかし、引用・参考文献のルールを全く遵守していないという致命的な形式的欠陥 (5 点) があり、学術レポートとしての信頼性が大きく損なわれています。内容の論理構成も、主張の核となる議論の深掘りが不足しているため、説得力が限定的です。形式面の改善と、議論の論拠をより学術的に固めることが、今後の必須課題です。

フィードバック文 (学生向け)

件名: レポート「E スポーツオリンピックについて」に関する評価と改善提案

レポート作成お疲れ様でした。e スポーツのオリンピック採用というテーマに対し、ニュース報道から世論の反応、そしてご自身の見解へと、流れを意識して構成した点は良くできています。特に、反対意見の根源として「年齢層による偏見」を指摘した視点はユニークで説得力があります。

レポートを合格点からさらに引き上げ、あなたの文章の「癖」を修正し、学術的な表現力を身につけるために、以下の点を次回必ず実践してください。

1. 学術的引用ルールの徹底 (最重要課題):

実践的な改善提案: あなたのレポートは、複数のニュースサイトや世論調査 (koeeru, 産経, 東洋経済など) の情報を基に構成されています。次回以降、本文中で外部の情報源から得た事実、意見、引用を記述した際は、その直後に必ず (著者名, 年) や (メディア名) といった具体的な出典をカッコ書きで明記してください。これを怠ると、あなたのレポートは「客

観的な根拠に乏しい」と評価されてしまいます。これは、学生が陥りがちな、情報を鵜呑みにするだけでなく、必ず出所を明示するという習慣づけに必要なステップです。

2. 文の区切り方と明確性の向上:

実践的な改善提案: あなたの文章は、複数の情報や判断を一つの長い文に詰め込みがちです。特に要約を行う際、接続詞や句読点が不足し、文意が不明瞭になっています。「一文一義」を徹底し、一つの文には一つの論点や事実のみを記述するように心がけてください。文を短く区切ることで、あなたの主張がより明確に伝わり、読みやすさが大きく向上します。

3. 議論の深掘り:

実践的な改善提案: 「マインドスポーツ」を引き合いに出して e スポーツへの批判に反論していますが、これは議論の深掘りというよりも比喩に留まっています。今回は、マインドスポーツが IOC に採用された具体的な背景や、e スポーツと伝統スポーツを分ける「スポーツの定義」の核となる要素を分析し、より学術的な観点から反論を展開してください。単に「他のものもある」というだけでなく、なぜ e スポーツがオリンピックの精神に合致しうるのか、積極的な論証が必要です。

4. 結論の強化:

実践的な改善提案: 結論部分 (第3章の終わり) は、レポート全体の主張を統合する場所です。あなたの最終結論は「年齢層上の人への理解、浸透が必要」となっていますが、これはレポートが解決すべき課題の一部です。序論で提示された問題提起 (スポーツの定義、依存症など) を踏まえ、あなたの主要な論点 (歴史、マインドスポーツ、偏見の克服) をすべて含んだ上で、「e スポーツのオリンピック採用は是である」という主張を力強く再確認し、締めくくってください。

学生 E へのフィードバック

総合評価: 60 点 (100 点満点)

その理由:

本レポートは、e スポーツの社会的な影響、特に認知度向上や経済的效果といった賛成理由と、健康問題や倫理的懸念といった反対理由を包括的に扱おうとしており、テーマに対する関心の高さは評価できます。しかし、学術的なレポートとして必須である「引用と参考文献」のルール遵守がほぼゼロであり (5

点)、主張や論拠の根拠を示すことができていません。また、文構造が複雑で読みにくいこと、自分の意見の章で議論の焦点がブレてしまい、最終的な結論が弱くなっている点も、論説文としての説得力を損なっています。アイデアは多角的ですが、それを適切に構成し、学術的な厳密さをもって表現するスキルが今後の課題です。

フィードバック文 (学生向け)

件名: レポート「e スポーツのオリンピック参加について」に関する採点結果と具体的な改善提案

レポート作成お疲れ様でした。e スポーツのオリンピック参加の是非というテーマに対し、賛否両論の意見を整理しつつ、ご自身の賛成の立場から具体的な意見(特にマイナーな競技選定への批判)を述べた点は評価できます。

しかし、このレポートをさらに高いレベルに引き上げ、学術的な文章作成スキル(文章の「癖」を修正する視点)を身につけるために、以下の点を次回必ず実践してください。

1. 学術的ルールの徹底遵守 (最重要課題):

実践的な改善提案: あなたのレポートには、外部の調査結果、ニュース、特定の事実など、多くの情報が使われています。これらは全て、誰が、いつ、どこで発表した情報かを明示しなければなりません。次回以降、本文中で外部情報を使用する際は、必ずその直後に(著者名、年)などの形式で出典を示してください。また、末尾の参考文献リストも、引用した情報全てを、正確かつ統一された形式(例: APA スタイルなど)で記載し直してください。情報源を明確にすることは、あなたのレポートの信頼性、つまり論証の土台となります。

2. 文の明確化 (一文一義の徹底):

実践的な改善提案: あなたの文章には、長い一文の中に複数の主語や結論が詰め込まれてしまう傾向があります。これにより、読者はどこがあなたの意見で、どこが引用した情報なのかを判断しにくくなっています。冗長な表現を避け、一つの文には一つの主要なメッセージや判断のみを含める「一文一義」の原則を徹底してください。文が長くなるときは、必ず句読点や接続詞で区切り、論理的な関係が明確になるよう修正しましょう。

3. 議論の一貫性と結論の強化:

実践的な改善提案: 第3章「自分意見」では、e スポーツの社会的懸念(プロゲーマーの進路問題)を提起し

ていますが、これは主要なテーマ(オリンピック参加への是非)からやや逸脱しています。論説文の最終目的は、自分の主張を説得力をもって導くことです。レポートの最後は、全体の議論(賛成・反対の整理)を踏まえ、あなたがなぜ「e スポーツのオリンピック参加に賛成するのか」という主張を最も強力に締めくくる段落にしてください。社会的な懸念を述べる場合も、それがオリンピック参加の是非とどう結びつくのか、明確な論理線で結びつけるようにしてください。

これらの改善点、特に引用の形式を厳守することで、あなたのレポートは内容面、形式面の両方で飛躍的に向上するでしょう。頑張ってください。

学生 F へのフィードバック

総合評価: 60 点 (100 点満点)

その理由:

本レポートは、e スポーツ大会の背景、賛否両論の整理、そして自身の意見の表明という流れで、論説文としての情報を網羅しようとしており、特にサウジアラビアの社会変革と e スポーツの関係性といった具体的な論点を取り上げている点は評価できます(内容面は比較的良好)。しかし、学術レポートとして必須である「引用と参考文献」のルール(本文中での出典明記)が全く守られていないという形式的欠陥(5点)が、信頼性を著しく損なっています。また、自身の主張が、本論で提示した倫理的な課題への反論を欠いているため、論理的な一貫性にも課題が残ります。

フィードバック文 (学生向け)

件名: レポート「第一回オリンピック e スポーツ大会について」に関する評価と具体的な改善提案

レポート作成お疲れ様でした。2025 年の第 1 回オリンピック e スポーツ大会という新しいテーマに対して、サウジアラビア側の詳細な背景(ゲーマー国王、大規模な大会実績)と IOC 側の目的(若年層の取り込み)を丁寧に調査し、整理した点は高く評価できます。また、支持意見だけでなく、サウジアラビアの人権問題といった倫理的な懸念にも言及している点も優れています。

このレポートをさらに高い評価(80 点以上)にするために、あなたの文章作成の「癖」と、形式面で必須の改善点を以下に指摘します。

1. 学術的ルールの徹底遵守（最重要課題）：

実践的な改善提案：あなたのレポートは、事実情報、数値、他者の主張など、多くの情報を外部から得ていますが、どの情報がどこから来たのかを示す「文中引用（インテキスト・サイテーション）」が完全に欠落しています。これが、あなたのレポートの信頼性を最も大きく損なう「癖」です。次回以降は、本文中の情報を使用した箇所すべてに対し、必ずその直後に（出典名、年）などの形式でカッコ書きの引用を示してください。参考文献リストを作成した努力を無駄にしないよう、このルールを徹底してください。

2. 文の明確化と冗長な表現の回避：

実践的な改善提案：あなたの文章は、複数の情報を一つの長い文に詰め込みすぎる傾向があります（例：賞金、種目、前回の額を一つの文にまとめている）。「一文一義」を意識し、文を短く区切ることで、読み手が内容をより正確に理解できるようになります。情報を羅列する際には、句読点ではなく「。そして」「。また」のように、文を区切ることを試みてください。

3. 議論の論理的な完結と結論の強化：

実践的な改善提案：2章であなたはサウジアラビア開催に伴う人権問題や倫理的な懸念という重大な論点を取り上げました。しかし、3章の自身の主張（賛成）では、この倫理的な懸念に対する具体的な反論や解決策が示されていません。結論として賛成の立場を取るならば、反対意見が提起した倫理的ジレンマに対し、なぜIOCの決定がそれでも正当化されるのか、という論理的な裏付けを必ず行ってください。また、「結論」という章を設け、レポート全体の主張を総括することを忘れないでください。

学生Gへのフィードバック

総合評価：76点（100点満点）

その理由：

本レポートは、OEG創設というテーマに対し、ニュースの背景理解、世論の分析、そして自身の見解を具体的に論じており、特に格闘ゲームのデバイス問題に関する詳細な分析（内容の示し方）は非常に優れています（18点）。しかし、大学レポートの基本として引用と参考文献のルール（文中引用）が守られていないこと（10点）が、スコアを大きく押し下げる要因となりました。また、第三章で議論が終わってしまい、レポートの総括を行う「結論」が欠けているため、全体構成に弱点が見られます。

フィードバック文（学生向け）

件名：レポート「オリンピックeスポーツゲームズについて」に関する評価と具体的な改善提案

レポート作成お疲れ様でした。今回のレポートは、IOCの新しい取り組みであるOEGについて、ニュースと一般の意見を丁寧に整理した上で、ご自身の具体的な見解を展開しており、内容面で大変優れています。特に、第三章でストリートファイター6を例に出し、デバイスの種類やボタン増設といった技術的な詳細に踏み込んで「全員の条件を同じにする難しさ」を論じた点は、プレイヤーならではの視点であり、議論に説得力を持たせています。

しかし、このレポートをさらに高い評価（80点以上）にするために、あなたの文章の「癖」と、形式面で必須の改善点を以下に指摘します。

1. 学術的ルールの徹底遵守（最重要課題）：

実践的な改善提案：あなたのレポートは、多くの事実情報（トーマス・バッハ会長のコメント、OESのルール、サウジアラビアの大会実績など）を外部の情報源から得ています。しかし、これらの情報が本文中のどこで使われたかを示す「文中引用（インテキスト・サイテーション）」が欠落しています。この「癖」を修正するため、次回以降は、外部情報を使用する際には必ず、引用文や事実の直後に（出典名、年）などの形式でカッコ書きを入れ、どの情報源を参照したのかを明示してください。参考文献リストを作成した努力を活かすためにも、このルールは厳守してください。

2. 構成の完成（結論の作成）：

実践的な改善提案：現在、レポートは「自分の考え」という議論の途中で終わってしまっています。論説文では、最後に必ず「結論」の章を設けてください。この結論では、あなたが本レポートで示した主なポイント（IOCへの感謝、ルールの複雑性、競技団体への権限委譲の提言）を改めて要約し、「OEGの成功に必要な要素」といった形で、レポート全体のメッセージを強く締めくくる必要があります。

3. 文体の改善と客観性の強化：

実践的な改善提案：第三章の冒頭で「IOCに対する感謝である」という感情的な要素から書き始める癖が見られます。学術レポートでは、まずは「IOCの決定は、eスポーツコミュニティに大きな期待をもたらした」のように、客観的な事実やコミュニティの反応から書き出し、その後で個人的な見解を述べる訓練をしてください。また、引用文を導入する際、「～～

とコメントを残した」のように、誰の発言かを明確に示す導入文を設けることで、文の明確さが向上します。

これらの改善点、特に引用の形式を修正し、結論を明

確にすることで、あなたの論理的で詳細な分析が、より説得力のある学術レポートとして評価されるでしょう。

Designing and Evaluating Prompts for Automated Report Grading Using Generative AI

KOMODA Yuchi MIYAZAKI Takashi

Abstract

This study examines prompt design and evaluation for automated report grading as an example of generative AI application in university education. We compared human grading with AI grading using Google NotebookLM on reports submitted by seven seminar students. Evaluation was conducted across five criteria: “Sentence and Vocabulary,” “Paragraph Structure,” “Overall Organization,” “Presentation of Content,” and “Quotations and References.” The results showed that the AI achieved a higher average score and exhibited a tendency to avoid low ratings. This is likely because the AI prioritizes superficial consistency and struggles to fully capture logical flaws or depth of content. On the other hand, the AI’s ability to grade with consistent criteria suggests its usefulness as an auxiliary tool for educational assessment. Future work should focus on prompt optimization and the development of multi-layered evaluation criteria.