

教育評価と ChatGPT

統計アプローチによる実証的研究

Educational Evaluation and ChatGPT

Empirical studies using statistical approaches

渡辺 直人 (Naoto Watanabe)¹

抄録

定性的評価物が増えている昨今、教育評価においては評価の曖昧性が課題となっている。発展著しい近年の AI テクノロジーを用いれば、この課題解決の手掛かりを得られるのではないだろうか。そのためにも、まずは基礎研究が必須である。本研究ではその一つとして、AI の教育評価能力を測定した。実施に関し、OpenAI がリリースしている ChatGPT (ver. 3.5) を用いて、報道レポート・記事100本を3回ずつ評価させた。得られた得点をそれぞれ「High」群、「Middle」群、「Low」群の3群に分け、フリードマン検定及びホルム法多重比較を行った。その結果、すべてのペアで有意な差が認められた ($p < .01$)。

With the recent increase in the number of qualitative evaluation items, ambiguity in evaluation has become an issue in the field of educational evaluation. We believe that the use of the rapidly developing AI technology of recent years may provide clues to solving this problem. To this end, basic research is essential. In this study, we measured the educational evaluation ability of AI. We used ChatGPT (ver. 3.5) released by OpenAI to evaluate 100 news reports three times each. The obtained scores were divided into three groups, "High," "Middle," and "Low," respectively, and Friedman test and Holm method multiple comparisons were conducted. The results showed that all pairs were significantly different ($p < .01$).

キーワード

ChatGPT, 教育評価, OpenAI

利益相反の有無

本研究に関連し、開示すべき COI 関係にある企業等はない。

1. 目的

近年、世界のテクノロジー業においては、人工知能を搭載した、対話可能な AIChat プログラムの開発が盛んである。様々な企業・研究者が参画し開発しているが、その中でも OpenAI が公開した「ChatGPT」(2022年11月30日公開)は特に有名であり、今では社会事象にもなっている。

OpenAI は、サム・アルトマン (Samuel H. Altman)、グレッグ・ブロックマン (Greg Brockman)、イリヤ・サツキヴァー (Ilya Sutskever)、ヴォイチェフ・ザレンバ (Wojciech Zaremba)、ジョン・シュルマン (John Schulman)、イーロン・マスク (Elon Reeve Musk) によつ

¹ 責任著者. 和歌山信愛女子短期大学 保育科 講師(専任)

〒640-0341 和歌山県和歌山市相坂702番地2 E-Mail:nwatanabe@shinai-u.ac.jp

て2015年に設立された¹⁾、人工知能技術を利用し、様々なツールを開発している団体である。GPT シリーズ、Jim、RoboSumo、Dactyl、Coodex、Neural MMO、Spinning Up、Clip、MuseNet、OpenAI Gym、TensorFlow Probability、DALL-E、API、Safety Gym 等、様々なツールが開発・発表されている。特にその中でも、ChatGPT の動力ともいえる GPT (Generative Pre-trained Transformer) シリーズは世界的に注目されている。

ChatGPT とは、AI によって動かされているチャットサービスである。換言すれば、AI との会話が可能なサービスである。偏った内容以外、様々な質問に対応することができ、かつ即座に返答される。恰も人間が対応したかのような、自然な受け答えが可能な点が主な特徴である。

この ChatGPT は、言語モデルを用いて言語生成するものである。言語モデルとは、言語を単語の出現確率でモデル化したものである。この言語モデルは、一般的には再学習によってモデルの強化を必要とするが、OpenAI の GPT-3 は、膨大なテキストデータを用いることにより、再学習を必要としない言語モデルを作り出しているという²⁾。ChatGPT は幅広いタスクを可能としており、質疑応答のみならず、詩や歌、エッセイ、短編小説、プログラムコードの生成も可能である³⁾。これらの処理は、ChatGPT が過去に学習した大量のテキストデータを基に生成されているという。

その動力となっている GPT シリーズであるが、これは2023年4月現在ではバージョン4まで存在している。まず、初版となった GPT-1 は2018年にリリースされ、つづく GPT-2 は翌年の2019年に公開されている。そして、2022年3月には GPT-3 が発表された。ChatGPT は無償版と有償版が存在しており、無償版では GPT-3.5 が用いられている。有償版 (月額20ドル) では GPT-4 を使用することが出来る (2023年4月現在)。

ChatGPT は、上述したように汎用性が極めて高いことから、あらゆる職種での活用が期待されている。実際に海外では裁判で使用された事例もある⁴⁾。わが国においても農林水産省⁵⁾、総務省⁶⁾ が ChatGPT を一部業務で活用するとしており、他にも広島県⁷⁾、横須賀市や戸田市でも活用の検証が行われているという⁸⁾。また、松野官房長官は、先進7か国首脳会議 (2023年 G7 広島サミット) において、ChatGPT のあり方を検討することを示している⁹⁾。

特に昨今の教育業界において多くの検討がなされており、業務の補助としての活用が期待されている。株式会社教育ネットでは『学校で AI を活用するための ChatGPT ガイド』を作成し、学校教職員向けに公開している¹⁰⁾。また、文部科学省も、学校・教職員向けに取扱い指針の作成することを公表している¹¹⁾。海外でも同様に議論が行われ、シンガポールでは教育現場での AI ツールを容認する方向を示している¹²⁾。

ここで「教育」という営みに関して概観していきたい。教育には年間・月間・週間、そして本時の指導計画が存在する。その指導計画の中には、必ず目的、目標・ねらいが設定され、それを達成すべく授業・指導を行う。そして、そのねらいに対してどの程度達成できたかは、必ず「評価」されなければならない。評価の種類は、文部科学省の「確かな学力」¹³⁾ を概観すると「目標に準拠した評価」(絶対評価)、「集団に準拠した評価」(相対評価)、「個人内評価」が紹介されている。また評価には「知識・技能」「思考・判断・表現」「主体的に学習に取り組む態度」の観点別で評価を行うことを示している¹⁴⁾。

文部科学省も「指導と評価の一体化」を掲げているように、教育において評価は欠くことができない営みである。そのような中で、宮国ら¹⁵⁾ は「近年では様々な分野において、「学ぶ力」や「興味・意欲・関心」といった様々な観点や能力、学習者のパフォーマンスなど、成果の定義や定量的な測定が容易でない観点等を評価すること」が求められていると述べている。さらに近年では「ルーブリックなどを用いた評価手法の開発や実践が行われている。このような評価方法は、評価者の主観的な判断要素が入り込む余地があるため、その活用には一定の注意が必要であるが、評価結果の信頼性に関する問題についての知見の蓄積は不十分なままである。」¹⁶⁾ と示している。加えて文部科学省の調査では、中高生の保護者の約40%が、評価に教師の主観が入っているのではないかと不安に感じているという¹⁷⁾。また、椿本・赤堀¹⁸⁾ は、先行研究を参考に、人間の評価には曖昧性という問題を孕んでいること¹⁹⁾ ²⁰⁾、そしてその曖昧性は、各評価物から受ける影響・系列効果によるものであること²¹⁾ ²²⁾ を示している。このように、定性的評価物の評価においては、評価者の主観が入り混じってしまう問題を孕んでいるといえる。

この問題に対し、AI テクノロジーの観点から解決の糸口をつかむことができるのではないだろうか。AI の補助を受けられれば、系列効果を防ぎ、評価基準を安定的に維持することが可能になると考える。しかし、ChatGPT はサービスの提供から1年も経っていないこともあり、先行研究は皆無に等しい。ChatGPT の機能面を評価・検討した学術的試みに関し、わが国における学術ジャーナルで、日本語で執筆された論文に限れば、2023年3月現在では存在しない。ChatGPT が著しい勢いで普及し、様々な用途で活用が見出されている昨今、早急に学術的検討が必要であるだろう。まずは多くの検討や実験を経て、課題を浮き彫りにさせることから着手させなければならない。

以上、教育の評価において ChatGPT はどの程度活用できるのであろうか。これらを明らかにするためには、妥当性、再現性の検討が必要である。本研究では、ChatGPT は教育の評価に活用できるか否か、その実験結果を報告する。特に本研究は第一報として、評価の再現性を明らかにする。

2. 方法

ChatGPT を使用し、報道レポート・記事の評価させた。評価は100%評価で行った。手続きとしては、無作為に選択した報道レポート・記事を100本用意した。質問文は「以下のレポートを100点満点で評価してください。」と記し、その文言に続いて記事の内容を添付した。そして、同じ報道レポート・記事を3度評価させた。2023年3月に実験を行った。

報道レポート・記事は、様々な報道社の記事を掲載している「Yahoo! ニュース」²³⁾ を参考にした。無作為に抽出したためジャンルは問わず、国際的な課題に関心を寄せているニュースも含めば、エンターテインメント性の高いレポートも含まれる。

ChatGPT は、AI の回答に対して、続けてレスポンスすることが可能となっている。しかし、同枠に続けて評価を行うと、前の質問・回答が次の評価に影響を及ぼし得ると考えたため、本研究では評価の度に新たな枠を設けて評価を行った。

1つの記事を3度評価させたのち、求められた点数を降順で3群に分けた。例えば、評価点が「90、85、80」であれば、最大値である90点は「High」群に、中間値である85点は「Middle」群に、最小値である80点は「Low」群に分けた。また「90、90、80」のように、続けて同じ点数が求められた場合は、「High」群、「Middle」群それぞれに90点を、そして「Low」群に80点を配点した。

分析方法は、これら3群に対し、フリードマン検定及び多重比較（ホルム法）を求めた。分析には清水²⁴⁾ の、Microsoft Office の Excel アドインツールである HAD (ver. 17) を用いた。

なお、2023年3月現在、無償で使用可能な ChatGPT のバージョンが3.5である。有償版ではバージョン4が使用できるが、本研究では一般社会で広く使用されると考えられるバージョン3.5で実験を行った。

3. 結果

(1) 要約統計量

はじめに、得られたデータの要約統計量を以下に示す。「High」「Middle」「Low」の3群に分けて示す(表1)。

「High」群では、平均値は89.2、標準偏差は6.03、最大値は100、最小値は68であった。

「Middle」群では、平均値は84.32、標準偏差は7.96、最大値は95、最小値は60であった。

「Low」群では、平均値は77.47、標準偏差は9.77、最大値は90、最小値は50であった。

なお、クロンバックの α 係数は0.85と十分な数値が得られた。

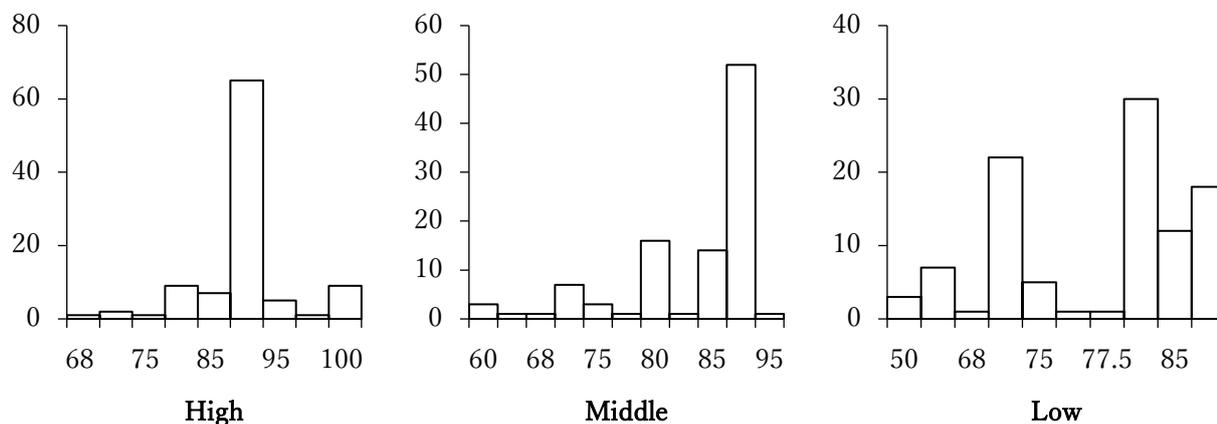
次に、度数(数値の出現回数)を「High」「Middle」「Low」の順に示す。

「High」群 100点:9回、97点:1回、95点:5回、90点:65回、85点:7回、80点:9回、75点:1回、70

点:2回、68点:1回 (表2)

「Middle」群 95点:1回、90点:52回、85点:14回、83点:1回、80点:16回、76点:1回、75点:3回、70点:7回、68点:1回、65点:1回、60点:3回 (表3)

「Low」群 90点:18回、85点:12回、80点:30回、77.5点:1回、76点:1回、75点:5回、70点:22回、68点:1回、60点:7回、50点:3回 (表4)



注) 縦線 : 度数、横線 : 点数

図1 各群のヒストグラム

表1 要約統計量

変数名	有効 N	平均値	中央値	標準偏差	分散	最小値	最大値
High	100	89.20	90	6.03	36.30	68	100
middle	100	84.32	90	7.96	63.41	60	95
Low	100	77.47	80	9.77	95.44	50	90

注) 100点満点で評価を行った。

表2 度数 (High)

出現値	度数	確率(%)	有効度数	有効確率	累積確率
68	1	1	1	1	1
70	2	2	2	2	3
75	1	1	1	1	4
80	9	9	9	9	13
85	7	7	7	7	20
90	65	65	65	65	85
95	5	5	5	5	90
97	1	1	1	1	91
100	9	9	9	9	100
欠損値	0	0			
合計	100	100	100	100	

表3 度数 (Middle)

出現値	度数	確率 (%)	有効度数	有効確率	累積確率
60	3	3	3	3	3
65	1	1	1	1	4
68	1	1	1	1	5
70	7	7	7	7	12
75	3	3	3	3	15
76	1	1	1	1	16
80	16	16	16	16	32
83	1	1	1	1	33
85	14	14	14	14	47
90	52	52	52	52	99
95	1	1	1	1	100
欠損値	0	0			
合計	100	100	100	100	

表4 度数 (Low)

出現値	度数	確率 (%)	有効度数	有効確率	累積確率
50	3	3	3	3	3
60	7	7	7	7	10
68	1	1	1	1	11
70	22	22	22	22	33
75	5	5	5	5	38
76	1	1	1	1	39
77.5	1	1	1	1	40
80	30	30	30	30	70
85	12	12	12	12	82
90	18	18	18	18	100
欠損値	0	0			
合計	100	100	100	100	

(2) フリードマン検定及び多重比較の結果

次に、フリードマン検定の結果を示す (表5)。「High」群、「Middle」群、「Low」群の3群を対象に検定を行った。平均順位は「High」群が2.71、「Middle」群が2.1、「Low」群が1.2であった。フリードマン検定の結果、有意な差が確認された ($\chi^2=152.3$, $df=2$, $\eta^2=51$, $p<.01$)。

次に、全てのペアをホルム法で多重比較を行った。その結果「High-Middle」($r=.31$, $p<.01$)、「High-Low」($r=.75$, $p<.01$)、「Middle-Low」($r=.45$, $p<.01$) の全てのペアで有意な差が確認された (表6)。

以上、今回行ったすべての分析において有意な差が認められた。ChatGPT (ver. 3.5) の評価が安定的であれば、差はでないことが仮説として成立する。同じ質問文で行っているため、乱数の存在・影響は考えにくい。だが、本結果ではその仮説は支持されず棄却された。このことから、ChatGPT (ver. 3.5) における評価は安定的であるといえないことが示された。

表5 フリードマン検定

検定統計量	152.30	$p < .01$
$\chi^2 =$		
$df =$	2	$\eta^2 = 0.51$

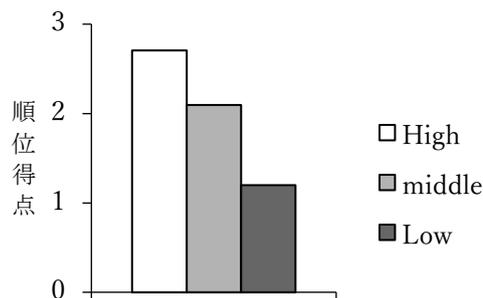


図2 得点順位のヒストグラム

表6 多重比較 (ホルムの方法)

ペア	順位差	効果量 r	95%CI	Z値
High - Middle***	.61	.31	.175, .425	4.31
High - Low***	1.51	.75	.686, .807	10.64
Middle - Low***	.90	.45	.330, .551	6.33

注) ***= $p < .01$

4. 考察

以上、本研究では ChatGPT (ver. 3.5) を使用し、無作為に選抜した報道レポート・記事100本を対象に、1本につき3度評価させた。3度の評価を「High」群、「Middle」群、「Low」群の3群に分け、フリードマン検定及び多重比較 (ホルム法) を行った。その結果、フリードマン検定で有意な結果が求められ、多重比較においても全ての群で有意な差が認められた ($p < .01$)。

評価に差があることから、現段階で ChatGPT (ver. 3.5) にレポート評価を行わせることは、適切とはいえないだろう。だが、仮に評価において ChatGPT を使用する状況がでてくるのであれば、一度の評価を採用するのではなく、複数回評価を行わせることは必須であろうし、また人間による確認も必要となるだろう。

また、ChatGPT は、質問の仕方で回答が変わることは指摘されているところである²⁵⁾。本研究では「以下のレポートを100点満点で評価してください。」という質問文を用いたが、質問の仕方次第で評価が変動する可能性は十分にある。加えて、この質問文は評価基準の自由度が高い。具体的に評価基準を定めておくことも一つの方法であろう。例えば、文章の評価基準においては、Remondino²⁶⁾ の作文評価「17観点」がある。そこには「①読みやすさ」「②美的配置」「③見かけ」「④綴り」「⑤単語の形式・語形変化」「⑥文法」「⑦アイデアの組み立て」「⑧アイデアの豊かさ」「⑨志向の適切さ」「⑩事実性」「⑪理解可能性」「⑫簡潔性」「⑬用語力」「⑭文体」「⑮独創性」「⑯完熟度」「⑰想像力」の17の観点が示されている (訳は渡部らによる²⁷⁾)。このような一定の基準を設けて実験を行えば、異なる結果が求められる可能性は否定できない。だが、この点は本研究では検討していないため、今後の課題として更なる考究が求められる。

加えて、本研究ではバージョン3.5を使用している。有償版であればバージョン4を使用でき、無論後者の方が高い精度を持っている。バージョンの違いで、結果が異なることも十分に考えられる。

ChatGPT は高い汎用性があるその反面、様々な問題を包含しているという。イタリアのデータ保護当局 (GDPD) は、ChatGPT には4つの問題点があると指摘している。OpenAI が子どもの使用における年齢管理をしていないこと、不正確な情報を提供する可能性があること、データ収集の対象者に、対象となったことが知らされていないこと。ChatGPT のトレーニングに使用される膨大なデー

タにおいて、人々の個人情報を収集する法的根拠がないことの4つであるという²⁸⁾。また、人工知能学会は、2023年4月に「まだ発展途上の技術であり、社会規範や倫理にそぐわないものを生成する可能性がある」と述べている²⁹⁾。他、社会においては既に多くの問題が報告されており、例えば ChatGPT に会社の機密情報を入力したという人為的な事故も一時話題となった。ChatGPT は AI の改良に使用すべく、入力されたデータは全て保存しているといい³⁰⁾、情報が拡散する恐れもある。加えて、一度保存されたデータの削除は不可能なため、OpenAI は機密情報の入力はしないよう警告しているという³¹⁾。

教育業界でも課題が浮き彫りになりつつあり、昨今では ChatGPT を用いた悪用が懸念されている。特に教育におけるレポート課題は、長らくこの問題に直面することになるだろう。既に大学のレポート課題を ChatGPT に代筆させた事例はある³²⁾。その事例では、2名の教授が AI による文章であることを発見したというが、後に「見抜くことは難しい」と述べている。今では、文章が AI によって生成されたものかを見抜く「AI チェッカー」等も開発されているが、ChatGPT や AI テクノロジーも進化している昨今では、これらチェックツールがどの程度対応できるかも未知である。

このように様々な問題が報告されている。これらのことから、ChatGPT の使用に制限をかける動きもみられ、例えばイタリアでは、一時的にはあるが使用禁止の措置を下している³³⁾。また、アメリカの事例では、ニューヨーク市教育局が学校のデバイス、ネットワークで ChatGPT をブロックしたという³⁴⁾。わが国においても、例えば鳥取県では業務上での ChatGPT の使用を禁止したことを発表している³⁵⁾。大学教育においても、例えば上智大学は、レポートや学位論文等での無許可な使用を禁止しており、東京大学では AI のみを用いたレポート作成を禁止するなど、様々な対策が施されつつある³⁶⁾。

以上、リリースされて間もない現在でも課題は多く、そして活用場面が広がるほど、今後も多くの問題が発生するだろう。この技術をどのように受け入れていくかは、慎重に議論されなければならない。

ただし、このようなサービスは絶対悪ではなく、これらをいかに活用していくか、その知見を見出し、開拓・発展させることこそ、今日の日本社会に求められていることなのではないだろうか。そしてこの機能を教育においてどのように活用するかは深い検討が必要であろうが、工夫次第で多くの道が開けるだろう。

最後に、本研究はわが国にほとんど存在しない ChatGPT における実証的研究であり、今後の研究発展に寄与するための基礎研究としての立ち位置である。本研究結果では ChatGPT (ver. 3.5) における評価の安定性は得られなかったが、上述したように AI テクノロジーは日々進化している。近年の AI 産業は目覚ましい発展を遂げており、近い将来この課題も解決される日がくるのではないだろうか。

文末脚注・参考文献

- ¹⁾ CNBC. OpenAI: 2023 CNBC Disruptor 50. <https://www.cnbc.com/2023/05/09/openai-disruptor-50.html>, 2023年4月20日取得.
- ²⁾ 野村総合研究所. GPT-3. https://www.nri.com/jp/knowledge/glossary/lst/alphabet/gpt_3, 2023年4月15日取得.
- ³⁾ GIZMODO. ChatGPT とは? 始め方や使い方、料金も紹介【4月11日更新】. <https://www.gizmodo.jp/2023/01/chat-gpt-openai-ai-finance-ai-everything-we-know.html#6>, 2023年4月20日取得.
- ⁴⁾ The Guardian. Colombian judge says he used ChatGPT in ruling. <https://www.theguardian.com/technology/2023/feb/03/colombia-judge-chatgpt-ruling>, 2023年4月21日取得.
- ⁵⁾ 日本経済新聞. 農林水産省、ChatGPT 活用へ マニュアル更新など. <https://www.nikkei.com/article/DGXZQOUA189IS0Y3A410C2000000/>, 2023年4月20日取得.

- 6) 産経新聞. 総務省もチャットGPT 松本総務相「試験的に利用」. <https://www.iza.ne.jp/article/20230420-V3T3I5AT3VJLJJHPG7ANPWUOH4/>, 2023年4月20日取得.
- 7) 日本経済新聞. 広島県、県行政にChatGPT活用検討 資料作成など. <https://www.nikkei.com/article/DGXZQOCC196CT0Z10C23A4000000/>, 2023年4月20日取得.
- 8) 日本経済新聞. 横須賀市や埼玉県戸田市、ChatGPTを実証 業務活用模索. <https://www.nikkei.com/article/DGXZQOCC1875R0Y3A410C2000000/>, 2023年4月20日取得.
- 9) 産経新聞. 広島サミットでチャットGPTの在り方討議 松野官房長官. <https://www.iza.ne.jp/article/20230420-SG762AYTFBNATBLTLCEBWMPU5Y/>, 2023年4月20日取得.
- 10) インプレス. 教育ネット、「学校でAIを活用するためのChatGPTガイド」を無料配布. <https://edu.watch.impress.co.jp/docs/news/1493895.html>, 2023年4月20日取得.
- 11) 産経新聞. チャットGPT 学習での取り扱い指針作成へ 文科省、5年度にも. <https://www.iza.ne.jp/article/20230406-UQUXRBKOTBNMRNZUQCR2HB2TGA/>, 2023年4月20日取得.
- 12) 株式会社知財図鑑. シンガポール政府、教育現場での「ChatGPT」などのAIツールの利用を容認. <https://chizaizukan.com/news/2BxpMjxKncSzzTtMuP4Ivx/>, 2023年4月20日取得.
- 13) 文部科学省. 確かな学力. https://www.mext.go.jp/a_menu/shotou/gakuryoku/faq.htm, 2023年4月10日取得.
- 14) 国立教育政策研究所. 学習評価の在り方ハンドブック 小・中学校編. https://www.nier.go.jp/kaihatsu/pdf/gakushuhyouka_R010613-01.pdf, 2023年4月10日取得.
- 15) 宮国泰史, 福本晃造, 杉尾幸司, 古川雅英 (2018). 教育評価における客観的評価と主観的評価との関係. 日本科学教育学会年会論文集, 42 (0), 263-266.
- 16) 同上.
- 17) 文部科学省. 児童生徒の学習評価の在り方について (報告). https://www.mext.go.jp/b_menu/shingi/chukyo/chukyo3/004/gaiyou/attach/1292216.htm, 2023年4月24日取得.
- 18) 椿本弥生, 赤堀侃司 (2007). 主観的レポート評価の系列効果を軽減するツールの開発と評価. 日本教育工学会論文誌, 30 (4), 275-282.
- 19) 平井洋子, 渡部洋 (1994). 小論文評定のカテゴリ化に関する推定論的考察. 行動計量学, 21, 21-31.
- 20) 渡部洋, 平由美子, 井上俊哉 (1989). 小論文評価データの解析. 東京大学教育学部紀要, 28, 143-164.
- 21) Cooper, P.L. (1984). The Assessment of Writing Ability: A Review of Research. Peter L. Cooper. GRE Board Research Report GREB No. 82-15R.
- 22) Klein, J. (2002). The Failure of decision support system: inconsistency in test grading by teachers. Teaching and Teacher Education, 18: 1023-1033.
- 23) Yahoo! Japan. Yahoo! ニュース. <https://news.yahoo.co.jp/>, 2023年3月1日取得.
- 24) 清水裕士 (2016). フリーの統計分析ソフトHAD: 機能の紹介と統計学習・教育, 研究実践における利用方法の提案. メディア・情報・コミュニケーション研究, 1, 59-73.
- 25) 吉田壘 (2023). 6 「ChatGPTを教育に取り入れる: アイデアと留意点」. <https://www.nii.ac.jp/event/other/decs/past.html>, 2023年4月20日取得.
- 26) Remondino, C. (1959). A factorial analysis of the evaluation of scholastic compositions in the mother tongue. British Journal of Educational Psychology, 30, 242-251.
- 27) 渡部, 平, 井上 (1989) 上掲.
- 28) WIRED. AIの学習データに含まれる個人情報、ChatGPTにとって“大問題”になる. <https://wired.jp/article/italy-ban-chatgpt-privacy-gdpr/>, 2023年4月26日取得.
- 29) 人工知能学会. 人工知能学会としての大規模生成モデルに対してのメッセージ. <https://www.ai-gakkai.or.jp/ai-elsi/archives/info/%E4%BA%BA%E5%B7%A5%E7%9F%A5%E8%83%BD%E5%AD%A6%E4%BC%9A%E3%81%A8%E3%81%97%E3%81%A6%E3%81%AE%E5%A4%A7%E8%A6%8F%E6%A8%A1%E7%94%9F%E6%88%90%E3%83%A2%E3%83%87%E3%83%AB%E3%81%AB%E5%AF%BE%E3%81%97%E3%81%A6>, 2023年4月26日取得.

- ³⁰⁾ GIZMODO. サムスン、機密情報を ChatGPT にリークして大問題に. <https://www.gizmodo.jp/2023/04/samsung-leaks-confidential-information-to-chatbot.html>, 2023年4月20日取得.
- ³¹⁾ 同上.
- ³²⁾ INSIDER. Two professors who say they caught students cheating on essays with ChatGPT explain why AI plagiarism can be hard to prove. <https://www.businessinsider.com/chatgpt-essays-college-cheating-professors-caught-students-ai-plagiarism-2023-1>, 2023年4月21日取得.
- ³³⁾ TV 朝日. 対話型 AI 「チャット GPT」 イタリアで一時使用禁止に. https://news.tv-asahi.co.jp/news_international/articles/000293752.html, 2023年4月20日取得.
- ³⁴⁾ Chalkbeat. NYC education department blocks ChatGPT on school devices, networks. <https://ny.chalkbeat.org/2023/1/3/23537987/nyc-schools-ban-chatgpt-writing-artificial-intelligence>, 2023年4月30日取得.
- ³⁵⁾ 朝日新聞デジタル. 鳥取県、業務では ChatGPT 禁止 知事「ちゃんとジーミーチー」. <https://www.asahi.com/articles/ASR4N4TS2R4NPUUB004.html>, 2023年4月21日取得.
- ³⁶⁾ 読売新聞オンライン. チャット G P T、学生の利用に対策…上智大「論文使用なら厳格な対応」. <https://www.yomiuri.co.jp/kyoiku/kyoiku/news/20230408-0YT1T50388/>, 2023年4月9日取得.