

AI (ChatGPT-4)、人間、および人間と AI の共同作成によって  
生成された宣伝文に対する人々の満足度

田中葵<sup>a</sup> 西山綾香<sup>b</sup> 中島一実<sup>c</sup> 本田翔<sup>d</sup>

要約

現在、生成的人工知能(AI)技術の一種であるChatGPTが話題になっている。一般的に、人間がタスクを行うよりもアルゴリズムがタスクを行う方がパフォーマンスが優れている場合が多い。にもかかわらず、タスクを実行する際、アルゴリズムに依存することを嫌うというアルゴリズム嫌悪の存在が明らかになっている。本研究は、生成AIの一種であるChatGPTに対しても、従来のアルゴリズム嫌悪と同様に、嫌悪感が見られるという仮説を基に分析を行った。仮説を支持する結果は得られなかったが、本研究の設定はあくまで一例に過ぎず、研究の設定次第では嫌悪感の強弱が変化すると考えられる。今後は、様々な状況設定で研究を行うことで、ChatGPTに対する嫌悪感の研究において示唆が得られると考える。

JEL分類番号: D90, D91

キーワード: ChatGPT, アルゴリズム嫌悪, キャンペーン促進, 再試

<sup>a</sup> 田中葵 同志社大学商学部 cgfg2100@mail3.doshisha.ac.jp

<sup>b</sup> 西山綾香 同志社大学商学部 cgfg0460@mail3.doshisha.ac.jp

<sup>c</sup> 中島一実 同志社大学商学部 cgfg0415@mail3.doshisha.ac.jp

<sup>d</sup> 本田翔 同志社大学商学部 cgfg0121@mail3.doshisha.ac.jp

## 1. イントロダクション

現在、生成的人工知能(AI)技術である支援チャットボットのChatGPTが話題になっている(Noy et al. 2023). 2022年11月末にChatGPTを公開してからわずか5日で利用者は100万人を達成した(日本経済新聞 23). ChatGPTは文章の作成や要約, 誤字脱字のチェック, さらに新たなアイデア創出に活用される(毎日新聞 23). 実際にChatGPTを用いて行われたオンライン実験では, 453人の大卒の専門職に対して, 職業に特化したインセンティブ付きのライティングタスクを割り当て, その半数にChatGPTをランダムに適用した. その結果ChatGPTは平均所要時間を40%減少, アウトプットの質を18%上昇させ, 生産性の大幅な向上に貢献した(Noy et al. 2023). このような実験結果から, 今後ChatGPTがより活用されていくと考える.

多くの先行研究では, 人間が行うタスクよりもChatGPTなどのアルゴリズムの方がパフォーマンスが優れている場合が多いにもかかわらず, 消費者は通常人間が行うタスクをアルゴリズムに依存して実行することを嫌うというアルゴリズム嫌悪(algorithm aversion)の存在が明らかになっている(Jussupow et al. 2020, Dietvorst et al. 2014). ところが, Zhang et al. (2023)では, AI(ここではChatGPTのことを指す)が関与して生成されたコンテンツを知っていても, コンテンツの知覚品質は低下せず, 結果としてアルゴリズム嫌悪は見られなかった.

ChatGPTは生成AIの一種であるため, 従来のアルゴリズム嫌悪と同様にChatGPTに対しても嫌悪感が見られると考える. そのため, アルゴリズム嫌悪が見られなかったZhang et al. (2023)の研究に興味を持ち, 再試を行うこととした. なお本研究では, 参加者の宣伝文に対する満足度を, 嫌悪感を測定するための指標として用いる. 本研究は, ChatGPTに対する嫌悪感の研究において示唆を与えると考える.

仮説: 従来のアルゴリズム嫌悪と同様に, ChatGPTに対しても嫌悪感が見られる.

## 2. サーベイ実験

### 2.1. 目的

日本のデータを用いた分析では海外のデータを用いた分析と異なる結果が見られることもあることから, Zhang et al. (2023)の研究の準レプリケーションを行い, 嫌悪感を調査した.

## 2.2. 実施期間・参加者数

実験実施期間は、2023年8月31日から2023年9月1日であり、最終的な有効回答数は10代から80代の1111名(男性738名, 女性369名, その他4名)となった。

## 2.3. 実験デザイン

本実験では参加者にシナリオを読んでもらい、作成者の異なるキャンペーンの宣伝文に対する満足度を調査した。

実験はクラウドソーシングサイトのYahoo!クラウドソーシングを用いて行った。本実験では、宣伝文を人間、ChatGPT-4、人間とChatGPT-4の共同作成の3つのパターンで作成し、それを参加者に提示した。次に、参加者を作成者に関する情報の提示の程度に応じベースライン条件、一部知っている条件、全て知っている条件の3条件に振り分けた。その後、各条件に参加者をランダムに割り当て、宣伝文の満足度を評価してもらった。参加者には、ベースライン条件では作成者について何も伝えず、一部知っている条件では作成者の種類について伝え、全て知っている条件では作成者の種類及び各宣伝文を作成したのがどの作成者なのかを伝えた。宣伝文はZhang et al. (2023)を参考に作成した。満足度評価後、ベースライン条件及び一部知っている条件の参加者には宣伝文の作成者を開示した上で、再度満足度の評価を行った。最後に個人属性について尋ねた。

## 2.4. 測定

実験中の満足度を測定する質問は「1:満足していない」から「6:満足している」までの6件法を用いた。また満足度評価後の、ベースライン条件及び一部知っている条件の参加者に宣伝文の作成者を開示した上で再度満足度の評価を行った際も、同様に6件法を用いた。

## 3. 結果

以下に記述統計(表1)、単純主効果(表2)、分散分析表(表3)を示す。

表1 記述統計

被験者内要因	被験者間要因	サンプルサイズ	満足度平均値	標準偏差
全て知っている条件	人間	384	4.1458	1.1263
	ChatGPT-4		4.2318	1.0625
	共同作成		4.4844	1.0790
一部知っている条件	人間	364	3.9560	1.1845
	ChatGPT-4		4.2005	1.0528
	共同作成		4.1731	1.0936
ベースライン条件	人間	363	4.0000	1.1467
	ChatGPT-4		4.1488	1.0511
	共同作成		4.1515	1.1571

表2 単純主効果

Source	Lambda	approx.Chi	df	p値	lower.bound	Greenhouse-Geisser	Huynh-Feldt-Lecoutre	Chi-Muller
全て知っている条件における全ての宣伝文	0.0608	5.5862	2	0.0612 +	0.5000	0.9857	0.9908	0.9907
一部知っている条件における全ての宣伝文	0.6503	0.8582	2	0.6511 ns	0.5000	0.9976	1.0031	1.0031
ベースライン条件における	0.0000	20.3455	2	0.0000 ***	0.5000	0.948	0.9529	0.9528

+p < .10, \*p < .05, \*\*p < .01, \*\*\*p < .001

表3 分散分析表

Source	SS	df	MS	F-ratio	p値
被験者間要因	25.1124	2	12.5562	4.5964	0.0103*
s×被験者間要因	3026.7981	1108	2.7318		
被験者内要因	32.1375	2	16.0688	34.1175	0.0000***
被験者間要因×被験者内要因	9.7718	4	2.4429	5.1869	0.0004***
s×被験者間要因×被験者内要因	1043.7002	2216	0.471		
合計	4137.9106	3332	1.2419		

+p < .10, \*p < .05, \*\*p < .01, \*\*\*p < .001

仮説を検証するため、ANOVAkun version 4.8.9を用いて参加者の各宣伝文に対する満足度について、2×3混合計画分散分析をおこなった。その結果、被験者内要因における相互作用のための単純主効果をみたところ、全て知っている条件の全ての宣伝文においてp=0.0612となり10%水準で有意差が見られた。また、分散分析表の被験者間要因×被験者内要因においてp=0.0004となり、1%水準で有意差が見られた。そのため、記述統計の結果より仮説は棄却された。

#### 4.考察

実験の結果、記述統計から、全て知っている条件の参加者における宣伝文への満足度は共同作成>ChatGPT>人間となっており、ChatGPTに対する嫌悪感が生じていなかった。むしろ、人とテクノロジーの共同作業が高く評価されていることや、ChatGPTに対する人以上の評価も観察された。そのため、我々の立てた仮説である「従来のアルゴリズム嫌悪と同様に、ChatGPTに対しても嫌悪感が見られる」は棄却された。

仮説が支持されなかった要因として考えられるのは、本研究における設定である。本研究の設定は、運動促進キャンペーンの宣伝文に対する満足度の評価であった。しかし、Castelo et al. (2019)では、主観的タスクではアルゴリズムが信頼されず依存度が低くなり、タスクの客観性を高めることでアルゴリズムへの信頼度と利用率が増加すると述べている。そのため、研究の設定次第では嫌悪感の強弱が変化すると考えられる。この点に関連して、たとえばLogg et al. (2019)は我々の実験結果と同様に、ある状況下では、人はアルゴリズムをより高く評価するという「algorithm appreciation」が生じることを実験で示している。我々の研究は、Logg et al. (2019)の結果と整合的である。

また、分散分析を行った結果から、ChatGPTと人間の共同作成が最も評価されていること、アルゴリズム嫌悪は起こっていないことが分かる。アルゴリズム嫌悪が見られなかった要因として、ChatGPTの流行によるバンドワゴン効果があると考えた。しかし、本研究の参加者の多くはChatGPTを使用したことがないと答えており、ChatGPTが流行しているため高く評価しているわけではなく、ChatGPTの能力自体を高く評価している可能性が考えられる。

## 5. 研究の限界と今後の展望

研究の限界は3点あげられる。まず1点目として、知っている条件の参加者に対して、作成者の意識付けが弱かった可能性がある。本研究では、参加者に3種類の作成者が作成した異なる宣伝文を見せ評価させた。しかし、参加者は作成者をあまり意識せずに文章を質のみで評価し、本研究での知っている条件の役割が果たされなかった可能性がある。

2点目は実験参加者の年齢層に偏りがあった点である。実験参加者の年齢層が高かったためかChatGPTを利用したことがない人の割合が多く、実験結果にも偏りが生じた可能性がある。

3点目は本研究で用いたテーマ設定が、あくまで運動の促進という1テーマにすぎなかった点である。テーマ設定によってはより嫌悪感が測定されやすいテーマもあり、本研究だけでは嫌悪感の

調査は十分でないと考えられる。以上のことから、今後の展望として、作成者の強調、年齢層の均一化、テーマ設定の多様化が必要である。本研究ではChatGPTを利用したことがない人の割合が多かったが、今後ChatGPTの利用率が上昇した場合、結果が変わる可能性がある。また、テーマ設定次第では嫌悪感の強弱が変わるため、様々なテーマでの実験の余地があると考ええる。

## 6.引用文献

Dietvorst, B. J., J.P. Simmons, and C. Massey (2014) Algorithm Aversion: People Erroneously Avoid Algorithms After Seeing Them Err. *Journal of Experimental Psychology: General*, 143 (6), 1-13.

E. Jussupow, I. Benbasat and A. Heinzl: Why are we averse towards algorithms? A comprehensive literature review on algorithm aversion, *Proceedings of the 28th European Conference on Information Systems(ECIS)*, 2020.

毎日新聞「チャットGPT:船橋市もAI試運用,効率化へチャットGPT」2023-07-20, 地方版/千葉, p.21.

Logg, J. M., Minson, J. A., & Moore, D. A. (2019). Algorithm appreciation: People prefer algorithmic to human judgment. *Organizational Behavior and Human Decision Processes*, 151, 90-103.

N. Castelo, M. W. Bos, and D. R. Lehmann, 2019. Task-dependent algorithm aversion. *Journal of Marketing Research*, 56(5) 809-825.

日本経済新聞「チャットGPT,数字で分析 1カ月の訪問者9億人,世界の1割 世界GDP押し上げ 効果は7兆ドル」2023-05-31

<https://www.nikkei.com/article/DGKKZO71477990Q3A530C2EA1000/>

S. Noy and W. Zhang, 2023. Experimental evidence on the productivity effects of generative artificial Intelligence.

高田, 2017. レプリケーション研究の方法. *JSMD Review*, 1(2), 65-71.

Y. Zhang, M. Sloan, B. Haas, R. Gosline and M. Sloan, 2023. Human favoritism, not ai aversion: people's perceptions (and bias) toward generative ai, human experts, and human-gai collaboration in persuasive content generation. SSRN.