

自己視点がSNS投稿における道徳的判断に及ぼす影響

Effects of the first-person perspective on moral judgement on SNS posting

加藤 仁

要旨

In this study, a preliminary survey and an online experiment were conducted with the followings: (1) to identify keywords co-occurred with flaming in a preliminary survey, (2) to clarify the effects of SNS posters' viewpoints and types of moral foundations on moral judgment, and (3) to propose intervention measures for self-censorship to prevent SNS flaming. The text data of posts on Twitter were collected and extracted words that co-occur with the word of "flaming". Based on these results and a database of the frequency of web search words, the keywords that predict flaming on SNS in terms of moral judgment were categorized. Then, Japanese ($n = 869$) participated in an online survey via a crowd-sourcing service. The results showed that posts were more likely to be evaluated as morally inappropriate when they contained moral violation words related to "harm". In addition, when the detection and warning of moral violation words were displayed on the screen before posting, it was found that the withdrawal of posts and the correction of keywords that may lead to SNS flaming were promoted.

キーワード：SNS 炎上 (SNS flaming)／道徳的判断 (moral judgement)／
自己中心性バイアス (egocentric bias)

I 問題と目的

ソーシャル・ネットワーキング・サービス (Social Networking Services; 以下, SNS) の利用者の増加は, オンラインの社会的ネットワークを拡大し, 人々が互いにアクセスすることを容易にした (総務省, 2013)。一方, インターネット上の「炎上」は2011年以降増加を続けており (総務省, 2019), SNSなどを通じて不特定多数の他者とコミュニケーションを取ることにリスクが伴う。

SNSにおける「炎上」とは, SNSなどのソーシャルメディアに個人や企業などの組織が投稿した内容が他の利用者の目に留まり, それに対する批判や反論を含むコメントが集中する現象をさす (i.e., flaming)。なお, インターネットにおける「炎上」に合意の得られた定義は存在していないが, 本研

究では田代・折田 (2012) にならい「情報発信者が管理するブログやSNS日記などの個人向けウェブサービスに嫌がらせコメントが殺到する現象」を炎上として定義し, 以降は炎上と表記する。

炎上には複数の種類が存在することが明らかになっており, 伊地知 (2007) は書き込まれる (投稿される) 内容によって「批判集中型」, 「議論過熱型」, 「荒らし」の3つのタイプに分類している。一方, 中川 (2010) では「義憤型」, 「いじめ・失望型」, 「便乗・祭り型」, 「不満・吐き散らし型」, 「嫉妬型」, 「頭をよく見せたい型」の書き込みの動機および内容によって6つのタイプに分類されている。また, 小林 (2011) は中川 (2010) のように, 炎上の対象となる事象によって「やらせ・捏造・自作自演」, 「なりすまし」, 「悪ノリ」, 「不良品・疑惑・不透明な対応」, 「コミュニティ習慣・規則の軽視」, 「放言・暴行・逆ギレ」の6つに分類している。このように投稿内容に着目するのか, またはその動機や対象に着目するのかで,

KATO, Jin

北陸学院大学 社会学部 社会学科
学習・言語心理学

研究者間でインターネット炎上の分類は異なっている。

SNS上ではその閲覧者を限定していない限り不特定多数の他者が投稿内容を閲覧することができ、投稿者の意図に関わらずSNS炎上が発生する可能性がある。SNS炎上のトピックとして、法に触れる可能性のある行為や他者から見て不快感のある行為は非道徳的であるとして非難されやすいと考えられる。道徳基盤理論 (Graham et al., 2011; Haidt, 2012) に基づくと、人は特に「加害」(i.e., 他者への暴力) や「不正」(i.e., 詐欺) に分類される意図・行為を目にした際にその非道徳性を認知し他者を非難しやすく、また非道徳的であると評価される他者は非難的になりやすい (Wegner & Gray, 2017)。

また、SNS炎上の背景として、投稿者が自らの投稿の道徳的な適切性に関する他者の反応を自己視点から推測して問題ないと判断し投稿する自己中心性バイアスに陥っている可能性が想定できる。自己中心性バイアスは自己視点の確証につながる結果、他者視点取得を妨げうる (Epley et al., 2004)。すなわち、SNS炎上は、SNS利用者が自己視点から投稿の是非を判断し、非道徳的であると認識されやすいトピックを含む投稿の不適切性を低く評価する結果として生じると考えられる。

そこで本研究では、「炎上」という単語と共起するトピックを特定し、SNS投稿者の視点および

道徳基盤に関連する投稿のトピックが道徳的判断に及ぼす影響を明らかにし、予防策を提案することを目的とし、予備調査・オンライン実験を実施した。

II 方法

予備調査

予備調査では、SNSで炎上しやすい投稿のトピックについて道徳性の5次元 (道徳基盤理論に含まれる道徳性の5次元: 「保護/加害」, 「公正/不正」, 「忠誠/不忠」, 「権威/不敬」, 「神聖/不純」) に基づいて分類を行った。2021年9-11月の3ヶ月間、Twitter上で公開されている投稿のテキストデータを収集し「炎上」という単語を含む466,277ツイート进行分析対象とした。このうち、重複ツイート、リプライ、ニュース、外部リンク紹介等を除くと286,649ツイートであった。続いて、RMacabで分かち書き、日本語の動詞と名詞に限定し、ストップワードおよび「炎上」という語を除外し、291,516語を抽出 (動詞: 15,689語、名詞: 275,827語) した (Figure 1)。最終的には、道徳基盤辞書日本語版 (Matsuo et al., 2019) を参照し各次元において出現頻度がもっとも高い単語を道徳違反語として抽出した (「保護/加害」次元: 被害, 「公正/不正」次元: 差別, 「忠誠/不忠」次元: 詐欺, 「権威/不敬」次元: 違反, 「神聖/不純」次元: 嫌悪)。それぞれの単語の出現

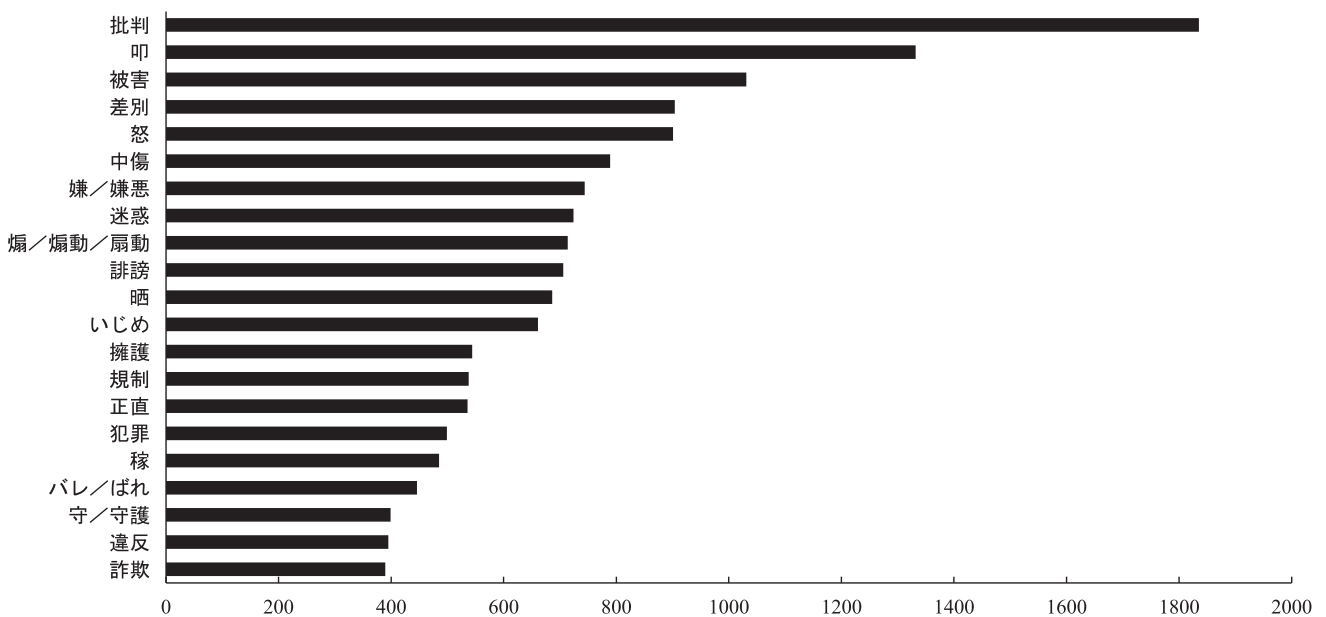


Figure 1. 「炎上」と共起するワードリスト (上位のみ抜粋)

回数は次の通りである；「保護／加害」次元：被害（1,031回出現）、「公正／不正」次元：差別（904回出現）、「忠誠／不忠」次元：詐欺（390回出現）、「権威／不敬」次元：違反（395回出現）、「神聖／不純」次元：嫌悪（744回出現）。これらの単語は後述する本調査の実験刺激（SNS上の投稿という形式）に含める形で条件操作（投稿のトピックの操作）を行うために使用した。本調査のために作成した道徳性の各次元に対応する実験刺激（SNSの投稿）はTable 1のとおりである（表内下線部が道徳違反語に該当）。

Table 1. 各道徳性次元のSNS投稿のトピック

「保護／加害」次元
「よくネットで話題になるこの手の事件って、被害者に対して賛否両論あるけど、個人的には違和感える…もう少し何とかならなかったのかなと思うけど。」
「公正／不正」次元
「よくネットで話題になるこの手の差別って、関係者に対して賛否両論あるけど、個人的には違和感える…もう少し何とかならなかったのかなと思うけど。」
「忠誠／不忠」次元
「よくネットで話題になるこの手の詐欺って、関係者に対して賛否両論あるけど、個人的には違和感える…もう少し何とかならなかったのかなと思うけど。」
「権威／不敬」次元
「よくネットで話題になるこの手の違反って、関係者に対して賛否両論あるけど、個人的には違和感える…もう少し何とかならなかったのかなと思うけど。」
「神聖／不純」次元
「よくネットで話題になるこの手の事件って、関係者に対して賛否両論あるけど、個人的には嫌悪感える…もう少し何とかならなかったのかなと思うけど。」
「中立」（統制条件）
「よくネットで話題になるこの手の事件って、関係者に対して賛否両論あるけど、個人的には違和感える…もう少し何とかならなかったのかなと思うけど。」

本調査

2021年12月、クラウドソーシングサービス（ランサーズ）を通じて1,000名を対象にオンライン実験を実施した（分析対象869名、男性497名、女性371名、平均39.32歳、 $SD = 9.75$ 歳）。

オンラインの質問紙実験では、道徳性の5次元（加害・不正・不忠・不敬・不純）に、道徳違反語を含まないニュートラルな投稿（中立）を加えた6つのトピックを作成し（Table 1）、各トピックに関するSNS上の投稿について、2つの異なる投稿者の視点（自己・第三者）から、投稿の不適切性（i.e., この投稿は、不適切である；5件法）と、投稿意図（i.e., 投稿者は実際に投稿すると思いますか；する・しないの2件法）を判断するよう求めた。また、実験参加者内要因として、全ての条件で投稿前に道徳違反語の検出・警告を画面上で強調表示した場合の投稿意図（i.e., 投稿者は実際に投稿すると思いますか；する・しないの2件法）と修正意図（i.e., 投稿者は投稿を修正すると思いますか；する・しないの2件法）について回答を求めた。最後に、次に示す各尺度項目についても回答を求めた。

質問紙構成

- （1）デモグラフィック項目：年齢、性別（女性、男性、その他）、婚姻状態（未婚、既婚、その他）、職業（会社員、自営業、経営者、公務員、学生、その他）について尋ねた。
- （2）SNS利用に関する質問：現在利用しているSNS（Twitter, Facebook, Instagram, TikTok, LinkedIn, Reddit）、SNS利用歴（10年以上、5-9年、2-4年、1年以下）、SNSの利用頻度（毎日、週6日、週5日、週4日、週3日、週2日、週1日、利用しない、その他）について尋ねた。
- （3）道徳基盤（金井, 2013）：「道徳原理への是認（MFQ2）」の16項目について「まったく同意しない」から「非常に同意する」の6段階で回答を求めた。
- （4）対人性反応性指標（IRI）（Davis, 1980）：「共感的配慮」、「視点取得」因子の各7項目について「まったく当てはまらない」から「とてもよく当てはまる」の5段階で回答を求めた。
- （5）批判的思考態度（平山・楠見, 2004）：「論理的思考への自覚」および「客観性」因子のうち、負荷量の高い各3項目について「あてはまらない」から「あてはまる」の5段階で回答を求めた。

(6) 社会的望ましさ(谷, 2008):「自己欺瞞」, 「印象操作」因子のうち, 負荷量の高い各3項目について「あてはまらない」から「あてはまる」の5段階で回答を求めた。

(7) ダークテトラッド:DTDD-J(田村他, 2015)に含まれる「マキャベリアニズム」, 「サイコパシー」, 「ナルシシズム」因子の各4項目と, 下司・小塩(2016)に含まれる「直接性サディズム」因子の4項目について「あてはまらない」から「あてはまる」の5段階で回答を求めた。

Ⅲ 結果

記述統計・相関分析

記述統計量および相関分析の結果をそれぞれTable 2およびTable 3に示す。また, 参加者のデモグラフィックについて, 現在利用しているSNSは, Twitter (86.31%), Facebook (37.63%), Instagram (56.50%), TikTok (8.75%), LinkedIn (2.42%), Reddit (0.46%)であった。SNS利用歴は, 1年以下(6.67%), 2-4年(23.94%), 5-9年(37.40%), 10年以上(31.99%)であった。SNS利用頻度は, 利用しない(0.70%), 週1日(9.80%), 週2日(6.30%), 週3日(8.05%), 週4日(3.85%), 週5日(6.53%), 週6日(2.33%), 毎日(62.43%)であった。

視点および道德性の次元がSNS投稿における道德的判断に及ぼす影響

投稿者の視点(参加者内要因)および投稿のトピック(参加者間要因)を独立変数, 投稿の不適

Table 2. 各変数の記述統計量

	Mean	SD	Min.	Max.
年齢	39.32	9.75	18.00	71.00
SNS利用歴	2.95	0.91	1.00	4.00
SNS利用頻度	5.46	2.23	0.00	7.00
道德基盤				
保護	13.86	2.39	6.00	18.00
公正	11.64	2.27	4.00	18.00
忠誠	11.23	2.25	3.00	18.00
権威	10.93	2.41	3.00	18.00
神聖	12.77	2.15	3.00	18.00
共感性				
共感的配慮	23.14	4.22	9.00	35.00
視点取得	24.20	3.96	12.00	35.00
批判的思考				
論理的思考への自覚	8.77	2.61	3.00	15.00
客観性	10.74	1.98	3.00	15.00
社会的望ましさ				
自己欺瞞	8.84	2.43	3.00	15.00
印象操作	8.44	2.23	3.00	15.00
ダークテトラッド				
マキャベリアニズム	8.75	3.37	4.00	20.00
サイコパシー	10.24	2.83	4.00	20.00
ナルシシズム	9.89	3.66	4.00	20.00
サディズム	6.76	2.66	4.00	18.00

Table 3. 各変数間の相関関係

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1. 年齢																	
2. SNS利用歴	-.102 **																
3. SNS利用頻度	-.156 **	.280 **															
道德基盤																	
4. 保護	.133 **	-.035	-.019														
5. 公正	.135 **	-.034	-.037	.361 **													
6. 忠誠	.099 **	-.065 +	-.071 *	.297 **	.193 **												
7. 権威	.003	-.040	-.047	.085 *	.000	.386 **											
8. 神聖	.064 +	-.031	-.077 *	.426 **	.294 **	.333 **	.245 **										
共感性																	
9. 共感的配慮	.109 **	.056 +	.026	.398 **	.315 **	.212 **	-.048	.249 **									
10. 視点取得	-.030	.047	.077 *	.265 **	.146 **	.193 **	-.006	.215 **	.473 **								
批判的思考																	
11. 論理的思考への自覚	.032	.047	-.010	.092 **	.057 +	.080 *	.059 +	.086 *	.213 **	.296 **							
12. 客観性	.000	.082 *	.069 *	.192 **	.184 **	.108 **	.013	.168 **	.347 **	.646 **	.405 **						
社会的望ましさ																	
13. 自己欺瞞	.037	.054	-.026	.046	.007	-.036	-.007	.031	.099 **	.098 **	.515 **	.187 **					
14. 印象操作	.180 **	-.110 **	-.042	.185 **	.179 **	.069 *	-.050	.166 **	.235 **	.113 **	-.016	.086 *	.009				
ダークテトラッド																	
15. マキャベリアニズム	-.208 **	.072 *	.009	-.223 **	-.145 **	-.063 +	.071 *	-.159 **	-.307 **	-.178 **	.009	-.183 **	-.025	-.604 **			
16. サイコパシー	-.186 **	-.002	-.025	-.324 **	-.194 **	-.198 **	.004	-.181 **	-.506 **	-.404 **	-.170 **	-.350 **	-.099 **	-.378 **	.557 **		
17. ナルシシズム	-.238 **	.095 **	.037	-.008	-.068 *	-.008	.038	-.055	.036	-.085 *	.059 +	-.118 **	.012	-.299 **	.426 **	.230 **	
18. サディズム	-.084 *	.008	-.046	-.261 **	-.120 **	-.097 **	.082 *	-.144 **	-.304 **	-.265 **	-.022	-.299 **	.022	-.377 **	.595 **	.536 **	.333 **

Note. ** $p < .01$, * $p < .05$, + $p < .10$

切性の判断を従属変数とする2要因分散分析を行った。分析の結果をFigure 2に示す。

その結果、投稿者の視点および投稿のトピックそれぞれの主効果がみられた(偏 $\eta^2_s = .022-.134$, $p < .01$)。投稿者の視点については、第三者ではなく自己による投稿の場合に、投稿の不適切性が高く評価されていた($d = -.149$, $p < .001$)。

投稿のトピックについて、Holm法による多重比較を行った結果、投稿に「保護／加害」次元の道德違反語が含まれる場合に、その他のトピックと比較して不適切であると評価されやすいことが明らかになった($ds = .181-.452$, $p < .01$)。

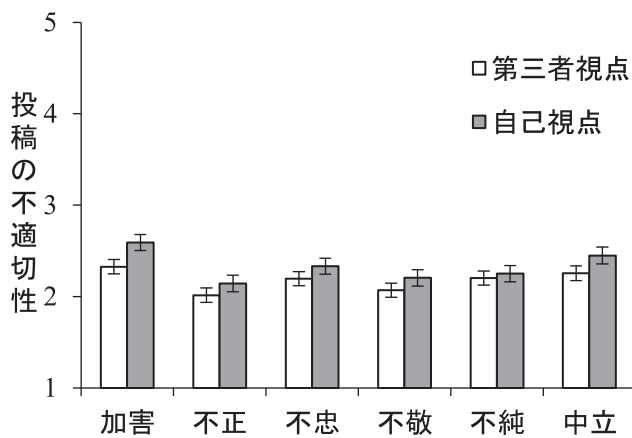


Figure 2. 視点および道德性の次元がSNS投稿における道徳的判断に及ぼす影響

視点が道德違反語を含むトピックのSNS投稿に及ぼす影響

投稿者の視点(自己視点または第三者視点)が道德違反語を含むトピックのSNS投稿を促進／抑制するかどうかを検討するために、視点間の投稿意図(投稿する、しない、わからない)と、道德違反語の検出・警告を画面上で強調表示した場合の投稿意図(修正オプションを付加)について χ^2 検定を行った。分析の結果をそれぞれFigure 3-5に示す。

その結果、第三者ではなく自己による投稿の場合に、「投稿しない」を選択しやすいことが明らかになった(Figure 3, 4; $V = .591$, $p < .01$)。さらに、道德違反語の検出・警告を画面上で強調表示した場合に、相対的に「投稿する」の割合が減少し、投稿の取り下げおよびSNS炎上の可能性のあるキーワードの修正が促進されることが明

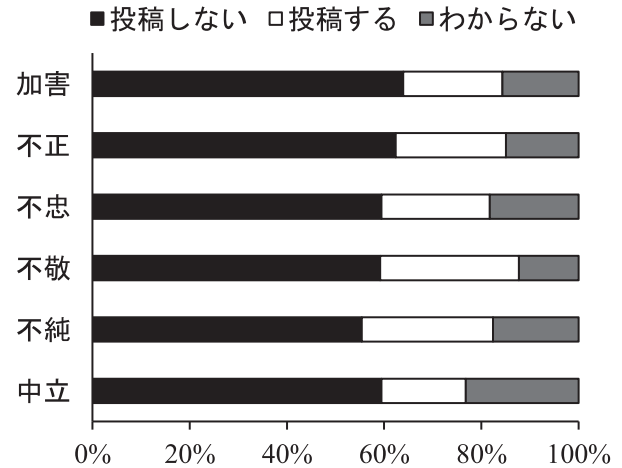


Figure 3. 自己視点における投稿意図

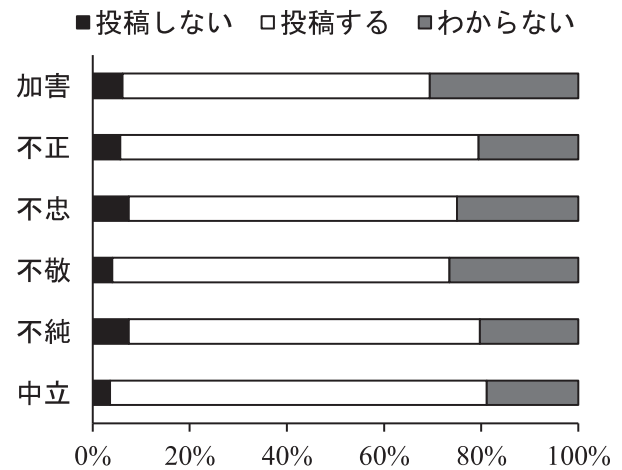


Figure 4. 第三者視点における投稿意図

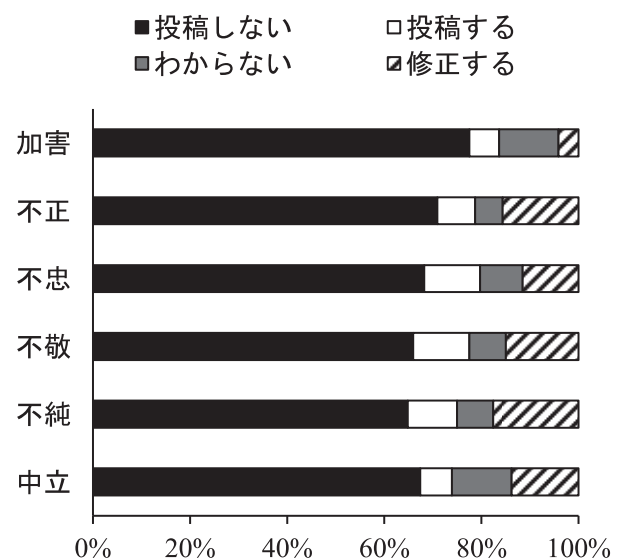


Figure 5. 警告を強調表示した場合の投稿意図

らかになった (Figure 5; $V_s = .213-.333$, $p < .01$)。

IV 考察

結果より、道德違反語が含まれる SNS 投稿が不適切であると判断されるのは「保護／加害」次元に限定されていた。加害行為を示唆する内容は特に非道德的であるとの判断を強める可能性が考えられる。一方、不適切性の判断の平均値は尺度の中点よりも低く、全体として不適切性が高く認識されていなかったことから、実験刺激として用いた SNS の投稿による操作が十分ではなかった可能性がある。

本研究では自己視点から道德的判断をすることで自己中心性バイアスが生じ、自己の正当性の確証に繋がると予測していたが、実際には自己の投稿の不適切性が高く評定された。SNS 投稿のように自己の責任や公的自己意識が顕現する場面では、投稿に対する慎重な態度が促進されと考えられる。また、リスク認知の観点からは自己のリスクを過大視するという方向で自己中心性バイアスが機能した可能性も想定できる。今後は自己中心性バイアスが道德違反語を含む投稿の促進／抑制のいずれの方向に働くかについて詳細に検討する必要がある。

SNS 炎上の予防策としては、炎上に繋がりうる道德違反語を検出したり、投稿内容の再考・修正を求めたりする警告表示システムを SNS に実装することで、SNS 投稿者の情報リテラシーの補助が可能になると考えられる。実際、本実験でも検出・警告によって投稿の再考が促進される可能性が示された。SNS への投稿時に事前の警告と修正オプションを設けるなど、投稿内容を再考・修正できる機会を提供することが SNS 炎上の回避につながる可能性がある。

限界と展望

本研究の限界として次の 2 点があげられる。第 1 に、実験参加者自身が SNS 投稿を作成していないことから、生態学的妥当性が十分に保証されていない可能性が考えられる。これを改善するために、実際に参加者が SNS 投稿を行う現実の場面での判断を求める必要がある。第 2 に、従属変数の測定において投稿の不適切性について尋ねる

こと自体が、自己の客観視を通じて他者視点取得を促進する結果、自己中心性バイアスによる自己の正当性の確証を抑制していた可能性が考えられる。この点は前述の自己中心性バイアスが機能する方向の問題であり、生態学的妥当性を担保する必要性和合わせて、実験手続きについて再考する必要がある。

〈引用文献〉

- Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *JSAS Catalog of Selected Documents in Psychology*, 10, 85.
- Epley, N., Keysar, B., Van Boven, L., & Gilovich, T. (2004). Perspective taking as egocentric anchoring and adjustment. *Journal of Personality and Social Psychology*, 87(3), 327-339.
- Graham, J., Nosek, B. A., Haidt, J., Iyer, R., Koleva, S., & Ditto, P. H. (2011). Mapping the moral domain. *Journal of Personality and Social Psychology*, 101(2), 366-385.
- Haidt, J. (2012). *The righteous mind: Why good people are divided by politics and religion*. Vintage.
- 平山 るみ・楠見 孝 (2004). 批判的思考態度が結論導出プロセスに及ぼす影響 証拠評価と結論生成課題を用いての検討 教育心理学研究, 52 (2), 186-198.
- 伊地知 晋一 (2007). ブログ炎上：Web2.0時代のリスクとチャンス アスキー
- 金井 良太 (2013). 脳に刻まれたモラルの起源—人はなぜ善を求めるのか— 岩波書店
- 小林 直樹 (2011). ソーシャルメディア炎上事件簿 日経 BP
- Matsuo, A., Sasahara, K., Taguchi, Y., & Karasawa, M. (2019). Development and validation of the Japanese moral foundations dictionary. *PloS one*, 14(3), e0213343.
- 中川 淳一郎 (2010). ウェブを炎上させるイタい人たち：面妖なネット原理主義者の「いなし方」 宝島社
- 下司忠大・小塩真司 (2016). 日本語版 Varieties of Sadistic Tendencies (VAST-J) の作成——因子構造および HEXACO との関連性—— 日本パーソナリティ心理学会第25回大会発表論文集, 109.
- 総務省 (2013). 情報通信白書 (平成25年版) ぎょうせい
- 総務省 (2019). 情報通信白書 (令和元年版) 日経印刷
- 田村 紋女・小塩 真司・田中 圭介・増井 啓太 (2015).

日本語版 Dark Triad Dirty Dozen (DTDD-J) 作成の試
み パーソナリティ研究, 24 (1), 26 - 37.

谷 伊織 (2008). バランス型社会的望ましさ反応尺度
日本語版 (BIDR-J) の作成と信頼性・妥当性の検討
パーソナリティ研究, 17 (1), 18 - 28.

田代 光輝・折田 明子 (2012). ネット炎上の発生過程
と収束過程に関する一考察—不具合に対する嫌がらせ
と決着による収束— 研究報告マルチメディア通信と
分散処理 (DPS), 2012 (6), 1 - 6.

Wegner, D. M., & Gray, K. (2017). *The mind club: Who thinks,
what feels, and why it matters*. Penguin

