

# SNS デザインの変更が誤情報識別に与える影響

## —「信頼ボタン」による介入の有効性—

HE YUZAN

SNS の普及により、誰もが容易に情報を発信・共有できるようになった。一方で、誤情報も拡散しやすくなり、深刻な社会問題となっている。近年の研究では、「いいね」などの肯定的なフィードバックが社会的インセンティブとして働き、情報の真偽を確かめないまま共有しやすくなる可能性が指摘されている。こうした背景のもと、Globig et al. (2023) は、「信頼する」「信頼しない」といった信頼性評価を可能にする反応ボタンを導入することで、真偽の識別や共有行動が改善されることを示した。しかし、この知見が日本においても再現されるかについては十分に検討されていない。そこで本研究では、日本人を対象に Globig et al. (2023) の研究 1 を追試し、「信頼する」「信頼しない」が「いいね」「よくないね」よりも正しい情報と誤情報の識別に有効であるかを検討する。

研究 1 では、日本人成人を対象に、Globig et al. (2023) の研究 1 を追試し、反応ボタンが情報の真偽判断に及ぼす影響を検討した。参加者は、正しい情報と誤情報からなる 30 件のニュース投稿を閲覧し、各投稿に対して「いいね」「よくないね」「信頼する」「信頼しない」の反応ボタンを用いて自由に反応した。その結果、「信頼する」「信頼しない」は、「いいね」「よくないね」よりも情報の真偽を識別できることを示した。一方で、肯定的反応と否定的反応の差は情報の種類(政治・科学・社会)によって異なる可能性も示唆された。

研究 2 では、研究 1 で確認された効果が特定の刺激内容に左右されないかを確認するため、一部の投稿を差し替えて追試を行った。その結果、反応の種類の主効果は研究 1 と同様に確認され、「信頼する」「信頼しない」は「いいね」「よくないね」よりも情報の真偽を識別できることを示した。一方で、肯定的反応と否定的反応の差は研究 1 とは異なるパターンを示し、識別力は刺激内容や提示形式といった文脈によって変動する可能性が示唆された。以上より、信頼性評価ボタンの効果は一定程度再現されたが、その大きさは情報の文脈に左右される可能性がある。

さらに、探索的に両研究で共通して用いられた投稿について研究間で比較したところ、真偽の識別は概ね同様の傾向を示した。ただし、刺激によって識別のされ方が変動する可能性も示唆された。

二つの実験を通じて、信頼性に関する反応ボタンが「いいね」「よくないね」に比べて、情報の真偽判断とより強く関連する可能性が示された。ただし、その効果は一様ではなく、情報の内容や文脈、反応の種類によって変動しうることも明らかとなった。したがって、反応ボタンへの介入は誤情報への有効な対策となり得るが、実装に当たっては文脈に応じた設計と運用が求められる。(社会心理学)