

流言に対する耐性獲得のための定期的な注意喚起手法の提案

西村 涼太^{1,a)} 平林 (宮部) 真衣^{2,b)} 吉野 孝^{1,c)}

概要: 近年、増え続ける流言に対し、様々なメディアや個人がその訂正情報を発信している。しかし、「訂正」による注意喚起は流言の発生に遅れを取ることが多い。また、膨大な数の流言に対して、それらを逐一確認して注意喚起するのは困難である。これに対して、人々をあらかじめ特定の言説に晒すことで、それらへの反論が可能な免疫を獲得する「接種理論」が注目されており、流言への応用が試みられている。一方で、接種効果が得られる方法については異論があり、確立されているとは言えない。また、その効果は徐々に減衰することが報告されている。そこで本研究では、我々の開発した流言の注意喚起システムによる効果の検証を行った。実験により、流言に関する定期的な注意喚起によって接種効果が発生し、注意喚起を直接行っていない流言へに対しても、ユーザの危機意識が向上することを明らかにした。また、実験結果から、流言に対する耐性を継続的に獲得することのできる手法について議論を行った。

Proposal of Periodic Alerting Method for Gaining Resistance to Rumors

1. はじめに

近年、SNSは情報の受信・発信の両面から重要なツールとなっている [1]。SNSでは誰もが情報を共有できるため、個人の情報発信が爆発的に増加している。それにとまって、フェイクニュースや流言*¹が増加し、アメリカの大統領選の混乱や COVID-19 のワクチン接種への反対運動を引き起こすなど、社会に大きな影響を与えている [2]。

これに対して、流言などの「注意を払うべき情報」を自動的に検出するための研究が多数行われている [3]。また、そうした技術を応用し、情報の閲覧者に対してリアルタイムに注意喚起を行うシステムの研究も行われている [4]。一方で、情報を「訂正」することで流言に対抗する手法は、逆にユーザの思い込みを深める場合もある [5]。また、個々の流言について注意喚起するのは、情報を受け取るユーザ

の負担を考慮すると量的な限界がある。加えて、多くの場合、訂正は流言の発生よりも後手に回るという課題もある。

そこで、最新の流言に関する研究では「接種理論」が着目されている [6]。接種理論は、生物医学的なウイルスに対するワクチン接種のアナロジーに基づき、これを情報に応用したものである [7]。すなわち、信じ込まない程度に弱められた言説への暴露による「接種」を通じて、説得への耐性を高めるのである。近年は、インターネットの普及とともにフェイクニュースや流言などに接種理論が応用され、その適用範囲が拡大している [8]。「接種」という用語は主に「ウイルス・ワクチンなどを人体や動物の体内に移植すること」を意味するものであるが、接種理論の流言への応用において、流言情報を人に暴露すると、人に情報に対する警戒心などを植え付けることに繋がると考えられる。そこで、本研究では情報へ暴露することを「接種」と呼ぶ。一方で、接種効果を得る手法については異論があり、また、その効果は徐々に減衰することが報告されている [9]。したがって、継続的な効果を得るには、追加の接種が必要となる [6]。しかし、既存のシステムは、特定のシナリオをこなす学習コンテンツ [10][11] など、一度きりの利用を想定したものが多く、継続的に効果を得る手法が確立されているとは言えない。

我々は、ユーザに流言の注意喚起を行うチャットボット

¹ 和歌山大学 システム工学部
Faculty of Systems Engineering, Wakayama University,
Wakayama 640-8510, Japan

² 東京大学大学院 医学系研究科
Graduate School of Medicine, The University of Tokyo,
Bunkyo, Tokyo 113-0033, Japan

a) nishimura.ryota@g.wakayama-u.jp

b) hirabayasi@m.u-tokyo.ac.jp

c) yoshino@wakayama-u.ac.jp

*¹ 本研究では、十分な根拠がなくその真偽が不明、または真偽が人々に疑われている情報を流言と定義し、発生過程での悪意の有無は問わないものとする。

「ちるも」を開発している。システムを通じて、ユーザはチャット形式で流言の情報を確認したり、注意喚起メッセージを定期的に受け取ったりすることができる。これまでの実験では、システムの有用性を確認し、流言に対してユーザの危機意識が向上する可能性を示した [12]。本稿ではちるもを介した、流言の注意喚起による効果の獲得について検証する。

2. 関連研究

2.1 接種理論

1960年代に McGuire が提唱した「接種理論 [7]」は、生物学のアナロジーに基づいている。弱められたウイルスを含む注射が抗体を作り、後の感染に対する抵抗力を高めるように、接種理論では情報でも同じことができると仮定している。すなわち、信じ込んでしまわない程度に弱められた条件下で人々を特定の言説に暴露させることで「接種効果」が発生し、後の説得への耐性を高められる。接種理論は、インターネットの普及とともにフェイクニュースや流言などに応用され、その適用範囲が拡大している [8]。最近の研究では、流言に共通する特徴や流言の拡散に用いられる手法に焦点を当てることで、特定の話題に依存しない汎用的な接種効果を得ることが試みられている [10]。一方で、接種効果は数週間かけて減衰することが観察されている [8][9]。また、一般的なワクチンと同様に、定期的な追加の接種を行うことで効果を延長できる可能性が示されている [13]。したがって、接種効果の持続期間に関わらず、定期的に接種を促す仕組みについて検討する必要がある。

2.2 接種理論を応用したシステム

接種理論を応用したシステムとして、Roosenbeek らの「Bad News」がある [10]。SNS においてフェイクニュースを生み出す過程を体験することで、ユーザはフェイクニュースを生み出す手法に関する接種を行うことができる。実験の結果、ゲームのプレイによってフェイクニュースを見分ける能力が向上することが示された。また、エコーチェンバー効果に対する接種を行うゲーム [11] など、いくつかの類似したシステムがある。このような、ゲーム型の接種システムでは、あらかじめ用意されたシナリオを達成する一度きりの使用が想定されている。また、ゲームプレイという能動的な行動が必要となるため比較的ユーザの負担が大きい。継続的に接種効果を得るには定期的な追加の接種が必要であり、その負担は小さいことが望ましい。そこで本研究では、能動的かつ一度きりの接種の手法でなく、受動的かつ定期的な接種が可能な仕組みについて検討する。

2.3 メディアリテラシー

近年の流言の爆発的な広がりを受けて、メディアリテラシー教育の必要性は増々高まっている。メディアリテ

ラシーについては、複数の特徴や能力が提唱されている。Tormala らによると、判断に対する自信は接種接種の過程において重要な役割を果たす [14]。また、Sanchez らによると、結論に早く飛びつく人ほど推論や決断で誤りを犯しやすく、陰謀論を信じやすい傾向がある [15]。また、Nagel らは、ニュースメディアリテラシーの測定項目「MAC2015」を提案している [16]。スコアが高い人ほどフェイクニュースを識別する自信があり、不確実な情報を共有する意欲が低く、ニュース消費に対して前向きで思慮深い行動が可能であると分析している。また、Maksl らが提案したメディアリテラシーの測定手法では、メディアリテラシーが高いほどニュース消費に対する内発的動機づけが強く、ニュースメディアに対して懐疑的であるとしている [17]。本研究ではメディアリテラシーに関する項目群を検証し、接種効果との関連性について考察する。

3. 流言の注意喚起システム「ちるも」

3.1 設計方針

「ちるも」は、Twitter 上から収集した流言にまつわる情報を用いて、ユーザに対して流言の注意喚起を行う。以下にシステムの設計方針を示す。

- (1) 流言に関してユーザに自己判断を促す 本システムでは、流言の真偽を断定せず、「誤りの可能性がある情報」としてユーザに提示する。情報の不確実性をユーザに認知させることを通じて、流言の真偽に関するユーザの主体的な判断を促進する。
- (2) 情報の真偽確認を容易にする 日々大量の情報が生み出される中で、逐一ファクトチェックを行う人は少数である [18]。本システムでは、ユーザは見聞きした情報をそのままチャット欄に入力するだけで、流言の情報を確認できるようにすることで、ユーザによる情報の真偽確認を容易にする。
- (3) ユーザの日常的な利用を促す 2.1 節で述べたように、接種効果は時間とともに減衰するため、定期的な追加の接種が必要となる。本システムは、日本で広く利用されているメッセージングアプリ「LINE^{*2}」上で動作するよう構築し、流言の注意喚起メッセージをユーザに定期的に通知する。これにより、ユーザが日常生活を送る中で、継続して接種効果を獲得できるようにする。

3.2 システムの基本機能

システムを中心となる画面例を図 1 (a) に示す。例では、ユーザの問いかけ「イギリスに関する流言を教えてください」に対して、関連するいくつかの流言（「イギリスは国葬を議会決議している」など）を返答している。また、マイペー

*2 <https://line.me>

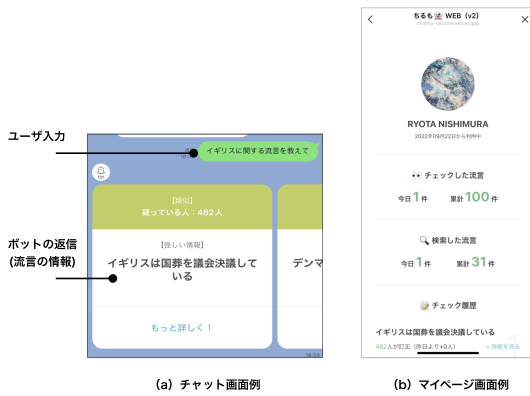


図 1 実験用のシステムのメインとマイページ画面



図 2 詳細と検索結果サンプル画面

ジ (図 1 (b)) では、これまでにシステム上で閲覧した流言の履歴など、ユーザの利用状況を確認できる。詳細ページ (図 2 (c)) では、「訂正ツイート*3」や、ちるも上での流言の定義をユーザに提示している。この画面で「Google 検索ボタン (図 2 (c))」を押すと、Google 社の検索エンジン*4を用いて流言について検索できる。検索を行う際のキーワードは、流言を形態素解析して得られた結果から名詞と動詞のみを抽出した単語群が自動的に入力される。図 2 (d) では、流言に関連するニュース記事が検索結果として表示されている。この機能によって、ユーザは手軽に追加の情報を調べることができる。各機能で利用する流言情報は、流言情報クラウド*5[19]と連携し取得している。

4. 実験

4.1 検証項目

流言の注意喚起と接種効果の獲得との関係を明らかにするため、本研究では以下の項目について検証する。

検証項目 1: 流言の注意喚起を行うことで、直接提示していない流言に対しても危機意識は向上するか？

*3 本研究では「流言を訂正する発言を含む Twitter 投稿」を訂正ツイートと定義する。

*4 <https://www.google.com>

*5 流言情報クラウドは、訂正ツイートから流言の情報を定期的に収集するシステムである。

検証項目 2: 流言の注意喚起を行うことで、メディアリテラシーに関する測定項目に変化があるか？

検証項目 3: 受動的な接種を中心とした仕組みで、継続的に接種効果を生じさせることは可能か？

4.2 実験概要

本研究では、システムを利用する実験群と利用しない対照群との比較実験を行う。参加者は、実験群が 30 名 (男性 16 名, 女性 14 名), 対照群が 26 名 (男性 14 名, 女性 12 名) で、なるべく男女比に偏りが無いよう無作為に割り振った。両群に共通して、実験開始前に事前アンケートを行い、実験開始から 7 日目に事後アンケートを行った。実験群では、事前アンケートの後、6 日間はシステムを自由に利用するように指示した。この実験期間中、毎日 20 時にシステムから流言の注意喚起メッセージを送信した。メッセージを送信する時間帯は、ユーザ負担の小ささと注意喚起効果の大きさを考慮して、以前の実験結果 [12] から決定した。対照群では、事前アンケートの回答後に、総務省が Web 上で公開している「ネットの時代におけるデマやフェイクニュース等の不確かな情報 [20]」を必ず 1 度は閲覧するよう指示した。この指示は、流言にまつわる実験そのものが参加者の危機意識に及ぼす影響を、できる限り実験群と同等にするためのものである。

4.3 事前・事後アンケート

本研究では、システム利用前後でのアンケートをオンライン上で行った。実験が終了するまでアンケートの内容をインターネットで検索したり他の人に話したりすることを禁止した。表 1 にアンケートの項目を示す。

質問 A から D は、流言を含む 12 項目の見出しを読み、それぞれの内容ごとに回答するものである。ただし、項目に流言が含まれていることは実験終了時まで参加者に知らせていない。そのため、実験終了後、参加者には流言の内容を知らせるデブリーフィングを実施した。また、実験前後での態度の変化を測定するため、事前・事後で用いている 12 項目の見出しは同一の内容とした。作成した見出しの詳細は 4.4 節で述べる。

質問 E から L は、ニュース消費に対する内発的動機づけと、ニュースメディアに対する懐疑心の強さを測定するものである [17]。

質問 M から P は、検証を補足するための質問群である。また、事前アンケートでは、これらの質問に加えて年齢や性別など社会的属性に関する質問も行った。

4.4 見出しの作成

本研究で使用する流言は、システムのデータベースから

表 1 事前・事後アンケートの質問項目

	質問	選択肢
A*	あなたは、上記の内容を知っていますか？	1：知っている，2：知らない，3：わからない
B	上記の内容を SNS 上で見たとして、あなたが共有する可能性はどの程度だと思いますか？（共有とは Twitter のリツイートなど、情報を拡散する行為）	1：非常に低い，2：低い，3：どちらともいえない，4：高い，5：非常に高い
C	上記の内容の「信頼性」はどの程度だと思いますか？	1：非常に低い，2：低い，3：どちらともいえない，4：高い，5：非常に高い
D	一つ前の「信頼性」に関するあなたの判断について、どの程度自信がありますか？	1：非常に自信がない，2：自信がない，3：どちらともいえない，4：自信がない，5：非常に自信がない
E	自分にとってニュースが何の役に立つかわからない	5 段階リッカートスケール（反転）
F	ニュースをチェックするのは、そうするのが当たり前だからだ	5 段階リッカートスケール
G	自分のために思って、ニュースをチェックしている	5 段階リッカートスケール
H	ニュースをチェックしているは、そうするのが好きだからだ	5 段階リッカートスケール
I	マスコミは公平だと思う	5 段階リッカートスケール（反転）
J	マスコミは一部始終を伝えていると思う	5 段階リッカートスケール（反転）
K	マスコミは、自分たちがいち早く報道することを最優先していると思う	5 段階リッカートスケール
L	マスコミの報道は正確だと思う	5 段階リッカートスケール（反転）
M**	実験期間を通じて、以前より流言（怪しい情報）を見分ける自信が向上した	5 段階リッカートスケール
N**	実験期間を通じて、見聞きした情報を調べて確認することが増えた（「ちるも」を利用した場合は、「ちるも」での検索・確認を含む）	5 段階リッカートスケール
O**	実験に参加する以前から、普段見ている情報の真偽には気をつけていた	5 段階リッカートスケール
P**	実験期間を通じて、見聞きした流言（怪しい情報）をいくつくらい覚えていますか？（「ちるも」を利用した場合は、「ちるも」で見たものを含む）	1：0 件，2：1～5 件，3：6～10 件，4：11～20 件，5：21 件以上

「5 段階リッカートスケール」は、1：強く同意しない，2：同意しない，3：どちらともいえない，4：同意する，5：強く同意する（反転）は逆転項目を示し、他の質問項目とは測定の向きが逆である

*事前アンケートのみで質問，** 事後アンケートのみで質問

抽出した 200 件の見出しに対し、特徴別*6に 6 つの区分に人手で分類したものである。分類区分からそれぞれ 1 件ずつ無作為に流言を抽出し、計 6 件をアンケートで用いた。また、真実のニュースを FIJ*7に参加している日本の主要なメディアの Web サイトから、流言と同数となる 6 件抽出した。真実のニュースは、Pennycook らの論文 [23] を参考に、政治、医療など 6 件の流言となるべく関連するカテゴリの話題から選定した。また、内容以外の差をできる限り減らすため、口語を含まない表現に統一するなど文章を調整した。事前・事後のアンケートでは、これら計 12 件の見出しを用いた。また、実験期間中にシステムから送信した注意喚起メッセージには、分類区分からそれぞれ 5 件ずつ無作為に抽出した流言を用いた。これらの流言は、事前・事後アンケートで用いた 6 件とは異なるものである。

5. 結果

5.1 信頼性

図 3 に信頼性評価（表 1 (C)）の平均値を示す。信頼性

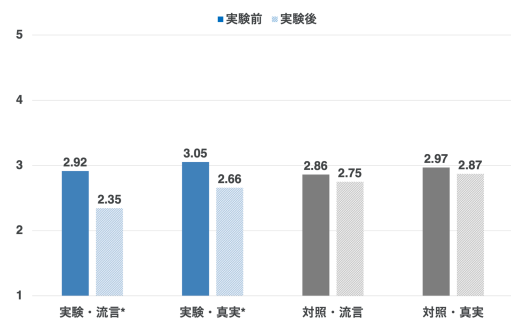


図 3 信頼性評価の平均値

* 有意差あり

評価の平均値は、参加者ごとの 12 項目の見出しに対する信頼性評価の平均値を計算し、さらに実験・対照群ごとの平均を算出したものである。この値について、実験群と対照群、流言と真実の見出しの 2×2 の組み合わせで検定を行った。検定はウィルコクソンの符号付順位検定を用い、ボンフェローニ補正 ($\alpha = 0.0125$) を行った。その結果、実験群では流言・真実の見出しとともに、実験後の信頼性評価が有意に低下していることがわかった（流言： $p = 0.0012$ ，真実： $p = 0.0009$ ）。また、実験群・流言の組み合わせで検

*6 分類は、NATO の発表している誤情報に関するレポート [21] および、EAVI (European Association for Viewers Interests) による「フェイクニュースの 10 分類 [22]」を参照した。

*7 日本でファクトチェックの普及活動を行う非営利団体

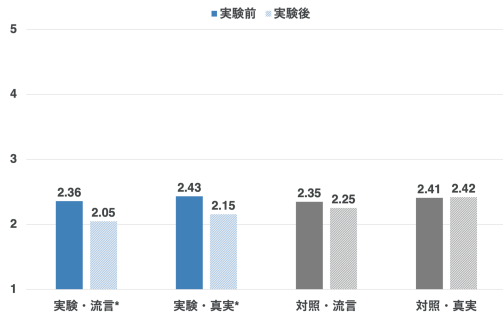


図 4 共有意欲の平均値
* 有意差あり

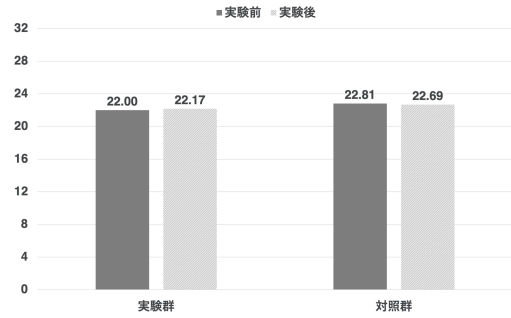


図 6 メディアリテラシースコアの平均値

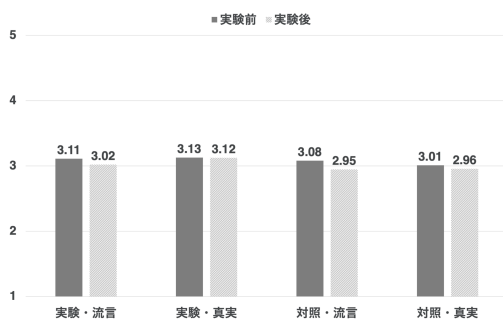


図 5 判断の自信度合いの平均値

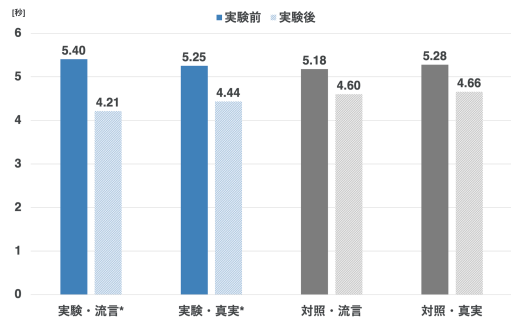


図 7 信頼性評価にかけた時間の平均値
* 有意差あり

定統計量 Z_t *⁸ および効果量 r *⁹を計算した。効果量 r は、 $0 \sim 1$ (または $-1 \sim 0$) の間で値を取り、関連性の強さを示す指標である。 $|r| > 0.5$ で大きい $|r| > 0.3$ で中程度、 $|r| > 0.1$ で小さい関連性があるとする。その結果、実験群・流言の組み合わせでは $Z = -3.24$, $r = 0.64$ となり、効果量が十分に大きいことがわかった。また、実験群・真実の組み合わせでは $Z = -3.33$, $r = 0.65$ となり、こちらも効果量が十分に大きいことを確認した。

5.2 共有意欲

図 4 に共有意欲評価 (表 1 (B)) の平均値を示す。計算方法および検定方法は 5.1 節と同様である。その結果、実験群では流言・真実の見出しともに、実験後の共有意欲評価が有意に低下していることがわかった (流言: $p = 0.0069$, 真実: $p = 0.0119$)。また、実験群・流言の組み合わせで検定統計量 Z_T および効果量 r を計算したところ、 $Z = -2.70$, $r = 0.55$ となり、十分に大きい効果量であることがわかった。実験群・真実の組み合わせでは、 $Z = -2.59$, $r = 0.50$ となり、こちらは中程度の効果量であることを確認した。

5.3 判断の自信度合い

図 5 に判断の自信度合い (表 1 (D)) の平均値を示す。計

*⁸ $Z_T = \frac{T - n(n-1)/4}{n(n+1)(2n+1)/24}$

*⁹ $r = \frac{Z}{\sqrt{N}}$

算方法および検定方法は 5.1 節と同様である。その結果、いずれの組み合わせでも実験前後で有意な差は見られなかった。また、それぞれの組み合わせで検定統計量 Z_T および効果量 r を計算したところ、実験群・流言では $Z = -0.32$, $r = 0.06$, 実験群・真実では $Z = -0.10$, $r = 0.02$, 対照群・流言では $Z = -1.17$, $r = 0.24$, 対照群・真実では $Z = -0.90$, $r = 0.18$ となり、いずれも効果量は小さいかほとんどないことがわかった。

5.4 メディアリテラシースコア

図 6 にメディアリテラシースコアの平均値を示す。メディアリテラシーの測定に用いた質問項目は表 1 の E から L の 8 項目である。メディアリテラシーは 5 段階リッカートスケールの 1 を 0 点、5 を 4 点として、0 点から 32 点の間で算出した。この値について、実験群と対照群でそれぞれ検定を行った。検定はウィルコクソンの符号付順位検定を用い、ボンフェローニ補正 ($\alpha = 0.025$) を行った。その結果、実験群・対照群ともに実験前後で有意な差は見られなかった。また、それぞれの組み合わせで検定統計量 Z_T および効果量 r を計算したところ、実験群では $Z = -0.08$, $r = 0.02$, 対照群では $Z = -0.38$, $r = 0.08$ となり、いずれも効果量は小さいかほとんどないことがわかった。

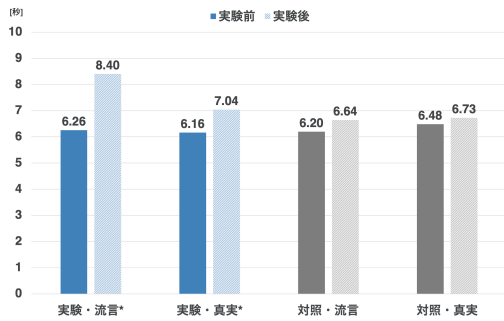


図 8 共有意欲評価にかけた時間の平均値
 * 有意差あり

5.5 評価にかけた時間

信頼性評価にかけた時間の平均値と、共有意欲評価にかけた時間の平均値を、図 7 と図 8 にそれぞれ示す。なお、データに極端に大きな値を取っているものが見られたため、IQR 法^{*10}を用いて外れ値として除外した。この外れ値は、数名の参加者が回答を一時中断したために発生したと考えられる。これらの値について、実験群と対照群、流言と真実の見出しの 2×2 の組み合わせで検定を行った。検定はウィルコクソンの符号付順位検定を用い、ボンフェローニ補正 ($\alpha = 0.0125$) を行った。その結果、信頼性評価にかけた時間の平均値は、実験群では流言・真実ともに有意に減少することがわかった (流言: $p = 0.0080$, 真実: $p = 0.0093$)。それぞれの組み合わせで検定統計量 Z_T および効果量 r を計算したところ、実験群・流言では $Z = -2.65$, $r = 0.52$, 実験群・真実では $Z = -2.60$, $r = 0.53$ となり、いずれも十分に大きい効果量であることがわかった。

一方で、共有意欲評価にかけた時間の平均値は、実験群では流言・真実ともに有意に増加することがわかった (流言: $p = 0.0080$, 真実: $p = 0.0072$)。それぞれの組み合わせで検定統計量 Z_T および効果量 r を計算したところ、実験群・流言では $Z = -2.65$, $r = 0.52$, 実験群・真実では $Z = -2.69$, $r = 0.60$ となり、いずれも十分に大きい効果量であることがわかった。

5.6 実験後アンケート

表 2 に実験後に行ったアンケート項目 (表 1 (M~P)) の結果を示す。質問 M 「実験期間を通じて、以前より流言 (怪しい情報) を見分ける自信が向上した」の結果について、ウィルコクソンの順位和検定により有意確率を求めると $p < 0.001$ であり、有意差が見られた。また、検定統計量 Z_T および効果量 r を計算したところ、 $Z = 5.06$, $r = 0.68$ であり、十分に大きい効果量であることがわかった。つまり、実験群では対照群に対して有意に流言を見分ける自信が向上したと回答した人が多いと考えられる。

*¹⁰ IQR (四分範囲) を用いて外れ値を検出する手法。

表 2 実験後アンケート

質問	グループ	評価の分布					中央値	最頻値
		1	2	3	4	5		
M	実験群	0	3	8	18	1	4	4
	対照群	4	13	8	1	0	2	2
N	実験群	0	2	3	21	4	4	4
	対照群	4	11	9	2	0	2	2
O	実験群	0	4	7	18	1	4	4
	対照群	0	3	8	14	1	4	4
P	実験群	1	19	7	3	0	2	2
	対照群	15	10	1	0	0	1	1

評価の分布: 1: 強く同意しない, 2: 同意しない, 3: どちらともいえない, 4: 同意する, 5: 強く同意する

表 3 ユーザー一人あたりの各機能の利用回数

機能	平均値 (回)	標準偏差 (回)	中央値 (回)
詳細を見る	7.0	5.9	6.0
マイページ	1.3	1.4	1.0
Google で検索する	2.2	2.3	2.0
ショートカット	2.5	2.0	2.0
話しかけて確認する	1.1	1.3	1.0

質問 N 「実験期間を通じて、見聞きした情報を調べて確認することが増えた (「ちるも」を利用した場合は、「ちるも」での検索・確認を含む)」の結果について、ウィルコクソンの順位和検定により有意確率を求めると $p < 0.001$ であり、有意差が見られた。また、検定統計量 Z_T および効果量 r を計算したところ、 $Z = 5.46$, $r = 0.73$ であり、十分に大きい効果量であることがわかった。つまり、実験群では対照群に対して有意に見聞きした情報を調べて確認することが増えたと回答した人が多いと考えられる。

質問 O 「実験に参加する以前から、普段見ている情報の真偽には気をつけていた」の結果について、ウィルコクソンの順位和検定により有意確率を求めると $p = 0.79$ であり、有意差は見られなかった。また、検定統計量 Z_T および効果量 r を計算したところ、 $Z = 0.27$, $r = 0.04$ であり、極めて小さい効果量であることがわかった。また、実験群・対照群ともに評価の中央値は 4、最頻値は 4 であることから、普段から情報の真偽には気をつけていると自認している参加者が多かったと考えられる。

質問 P 「実験期間を通じて、見聞きした流言 (怪しい情報) をいくつくらい覚えていますか? (「ちるも」を利用した場合は、「ちるも」で見たものを含む)」の結果について述べる。実験群の評価は中央値が 2、最頻値が 2 であることから、実験群では 1~5 件程度の流言を覚えていると回答した人が多いと考えられる。一方で、対照群の評価は中央値が 1、最頻値が 1 であった。また、対照群では、26 人中 15 人が流言を一つも覚えていないという結果になった。

5.7 利用ログ

表 3 に実験期間中のシステムの利用状況を示す。最も多く利用されたのは「詳細を見る」機能で、一人あたり平均で 7 回利用された。「ショートカット」と「Google で検索する機能」は、それぞれ一人あたり平均で 2.5 回、2.3 回利用された。「マイページ」と「話しかけて確認する」機能はこれらの中で最も少なく、それぞれ一人あたり平均で 1.3 回、1.1 回利用された。

6. 考察

6.1 接種効果の発生

本節では「検証項目 1：流言の注意喚起を行うことで、直接提示していない流言に対しても危機意識は向上するか？」を検証する。5.1 節で述べたように、実験群では実験前に比べて実験後の信頼性評価が有意に低下していることがわかった。ユーザが流言を含む情報をどの程度信頼するかは、接種効果を測定するために用いられる最も基本的な指標である [10][11][23]。接種を受けた参加者は、流言の信頼性に対してより低い評価を下す、すなわち、より懐疑的になることが示されている。この点において、本研究では先行研究と同様の結果が得られた。また、本実験では真実の見出しに関しても、流言と同様に信頼性評価が低下するという結果となった。この理由として、実験に用いた流言と真実の見出しの真偽を判断するのが困難だった可能性が考えられる。真実であることが明白な見出しを除外することは、評価実験において重要であり [23]、本実験ではこの点を満たしていると考えられる。また、実験に用いた真実の見出しは事前にファクトチェックを行っているが、現実のニュースではその真偽が後から覆ることもある。したがって、接種によって、ユーザの危機意識が流言・真実ともに一定程度まで向上した本研究の結果は、肯定的に捉えられる。これらのことから、ユーザは流言の注意喚起メッセージを受け取ったり、流言について調べたりすることを通じて、内容が同一でない流言に対しても危機意識を高めることができたと言える。したがって、提案手法によってユーザは接種効果を獲得できる可能性を示した。

6.2 メディアリテラシーへの影響

本節では「検証項目 2：流言の注意喚起を行うことで、メディアリテラシーに関する測定項目に変化があるか？」について検証する。5.2 節で述べたように、実験群では実験前に比べて実験後の共有意欲評価が有意に低下していることがわかった。不確実な情報を共有する意欲が低いことはメディアリテラシーの一つとして重要である [16]。また、本実験では真実の見出しに関しても共有意欲が低下した。実験期間中、参加者はアンケート項目に関して調べることが禁止されていたため、本実験の状況下で真実の見出しに関しても、不確かな情報は共有しないと判断を下すように

変化したのは肯定的な結果である。

次に、判断の自信度合いでは実験前後で有意な差が見られなかった。この理由として、前節で述べたように、実験に用いた流言と真実の見出しの真偽を判断するのが困難だったことが考えられる。正しい判断に対して自信が伴うことはメディアリテラシーにおいて重要である [24]。一方で、真偽の判断が困難な状況において、それを過度に自信を持って断定しないことも肯定的に捉えられる。

次に、メディアリテラシースコアでは実験前後で有意な差が見られなかった。本実験で測定したスコアは、実験群・対照群ともに実験前の時点で 22 点から 23 点である。これは最大の 7 割程度のスコアと比較的高いスコアである。また、実験前アンケート (5.6 節) でも、普段から情報の真偽には気をつけていると考えている参加者が多い傾向にあった。これらのことから、本実験ではスコアの有意な向上が見られなかった可能性がある。一方で、ある程度の注意力を持つと考えている人に対しても、システムによる接種効果があったと解釈することも可能である。

最後に、信頼性評価にかけた時間について、実験群では流言・真実ともに有意に減少することがわかった (図 7, 8)。一般的には直感的に素早く下す判断は、熟慮を伴う判断に比べて誤ることが多いと言われている [15]。しかし、6.1 節で述べたように、実験後にユーザはより懐疑的な判断を下すように変化しており、必ずしも短絡的に判断しているとは言えない。このことから、接種により、ユーザは短時間のうちに、より懐疑的な判断を下すよう変化していると考えられる。一方で、共有意欲評価にかけた時間について、実験群では流言・真実ともに有意に増加することがわかった。特に流言では 1 問につき一人あたり平均で 2 秒近く回答時間が増加している。これら結果より、信頼性、共有意欲の評価が実験群で有意に低下していることから、実験後のユーザは情報に対して疑いを持ち、情報共有には慎重な判断を下すように変化したと考えられる。

6.3 受動的なシステム利用の効果

本節では、「検証項目 3：受動的な接種を中心とした仕組みで、継続的に接種効果を発生させることは可能か？」について検証する。実験期間中には、一人あたり平均 7 回の詳細情報の閲覧があり、システム内の検索機能も平均 2 回以上利用されていた (表 3)。また、実験後のアンケート「実験期間を通じて、見聞きした情報を調べて確認することが増えた (表 2 (N))」では、中央値、最頻値ともに 4 (同意する) と回答した人が最も多かった。これらの結果から、設計方針の通り、注意喚起メッセージを閲覧する受動的な利用が主な利用形態となっていることがわかる。また、情報の閲覧に付随して、検索行動などの主体的な真偽確認行動を促進していることも考えられる。

既存研究では、学習型コンテンツの使用前後での比較な

ど、数十分間から数時間の間での比較が多く、長期的な利用による効果の検証は少ない。これに対し、本実験では、システム利用を含む約1週間の中で接種効果が発生し持続することを示した。すなわち、受動的な接種を中心とした仕組みによって、ユーザは接種効果を継続的に得られる可能性がある。これらのことから、提案手法はユーザにとって利用の負担が少なく、効果が減衰するという接種理論の問題に対処できる可能性がある。

7. 終わりに

本研究では、流言の定期的な注意喚起による接種効果を検証することを目的に、評価実験を行った。その結果、流言の注意喚起を受け取ることで接種効果が発生し、流言に対する耐性が向上することが明らかになった。また、提案手法は継続性に優れているため、接種効果が減衰する課題に対処できる可能性を示した。今後はより長期的なシステム利用による接種効果の発生について検証し、メディアリテラシーにまつわる各項目間の関係性についても検証する。

参考文献

- [1] Kristina Lerman, Rumi Ghosh: Information Contagion: an Empirical Study of the Spread of News on Digg and Twitter Social Networks, Proceedings of 4th International Conference on Weblogs and Social Media (ICWSM-10), 2010.
- [2] The far-right disinformation pipeline that now spans the globe goes straight from Trumpland to Australia, The Guardian (2022), <https://www.theguardian.com/commentisfree/2022/aug/18/the-far-right-disinformation-pipeline-that-now-spans-the-globe-goes-straight-from-trumpland-to-australia>, (最終閲覧: 2022-10-22).
- [3] Gupta, A. and Kumaraguru, P.: Credibility Ranking of Tweets during High Impact Events, Proceedings of the 1st Workshop on Privacy and Security in Online Social Media, No. 2, ACM, pp. 1-8 (2012).
- [4] Rob Ennals, Beth Trushkowsky, John Mark Agosta: Highlighting disputed claims on the web, Proceedings of the 19th international conference on World wide web April 2010 pp.341-350 (2010).
- [5] Mingkun Gao, Ziang Xiao, Karrie G Karahalios, Wai Tat Fu: To Label or Not to Label: The Effect of Stance and Credibility Labels on Readers' Selection and Perception of News Articles, Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'11), ACM, pp.1245-1254 (2011).
- [6] Stephan Lewandowsky, Sander van der Linden: Countering Misinformation and Fake News Through Inoculation and Prebunking, European Review of Social Psychology, Volume.32, Issue.2, pp.348-384 (2021).
- [7] McGuire, WJ, Papageorgis D: Effectiveness of forewarning in developing resistance to persuasion, Public opinion quarterly, 26(1), pp.24-34 (1962).
- [8] John A. Banas, Stephen A. Rains: A Meta-Analysis of Research on Inoculation Theory, Communication Monographs, Volume.77, Issue.3, pp.281-311 (2010).
- [9] Thomas Zerback, Florian Töpfl, Maria KnöpfleView: The disconcerting potential of online disinformation: Persuasive effects of astroturfing comments and three strategies for inoculation against them, New Media & Society, Volume.23, Issue.5, pp.1-43 (2021).
- [10] Jon Roozenbeek, Sander van der Linden: Fake news game confers psychological resistance against online misinformation, Palgrave Communications, volume.5, pp.1-10 (2019).
- [11] Youngseung Jeon, Bogoan Kim: ChamberBreaker: Mitigating the Echo Chamber Effect and Supporting Information Hygiene through a Gamified Inoculation System, Proceedings of the ACM on Human-Computer Interaction, Volume.5, pp.1-26 (2021).
- [12] Ryota Nishimura, Mai Miyabe Hirabayashi, Takashi Yoshino: Baseless-Rumor Alert Bot to Promote Reliability of Information, Journal of Information Processing, Volume.30, pp.2-14 (2022).
- [13] Rakoen Maertens, Frederik Anseel, Sander van der Linden: Combatting climate change misinformation: Evidence for longevity of inoculation and consensus messaging effects, Journal of Environmental Psychology, Volume.70, pp.1-11 (2020).
- [14] Zakary L.Tormala, Richard E.Petty: Source Credibility and Attitude Certainty: A Metacognitive Analysis of Resistance to Persuasion, Journal of Consumer Psychology, Volume.14, Issue.4, pp.427-442 (2004).
- [15] Carmen Sanchez, David Dunning: Leap of Confusion, Scientific American, February, pp.64-67 (2022).
- [16] Tyler W. S. Nagel: Measuring Fake News Acumen Using a News Media Literacy Instrument, Journal of Media Literacy Education, Volume.14, Issue.1, pp.29-42 (2022).
- [17] Adam Maksl, Seth Ashley: Measuring News Media Literacy, Journal of Media Literacy Education, Volume.6, Issue.3, pp.29-45 (2015).
- [18] Luping Wang, Susan R. Fussell: More Than a Click: Exploring College Students' Decision-Making Processes in Online News Sharing, Proceedings of the ACM on Human-Computer Interaction, Vol.4, No.9, pp.-20 (2020).
- [19] 宮部 真衣, 灘本 明代, 荒牧 英治: 人間による訂正情報に着目した流言拡散防止サービスの構築, 情報処理学会論文誌, Vol.55, No.1, pp.563-573 (2014) .
- [20] ネットの時代におけるデマやフェイクニュース等の不確かな情報, 総務省 (2022), https://www.soumu.go.jp/use_the_internet_wisely/special/fakenews/, (最終閲覧: 2022-10-22).
- [21] Digital hydra: security implications of false information online, NATO StratCom (2017), <https://www.stratcomcoe.org/digital-hydra-security-implications-false-information-online> (最終閲覧: 2022-10-21).
- [22] Beyond "fake news" - 10 types of misleading news, EAVI (2018), <https://eavi.eu/beyond-fake-news-10-types-misleading-info/>, (最終閲覧: 2022-10-21).
- [23] Gordon Pennycook, Jabin Binnendyk, Christie Newton, David G. Rand: A Practical Guide to Doing Behavioral Research on Fake News and Misinformation, Collabra: Psychology, Volume.7, Issue.1, pp.1-13 (2021).
- [24] Melisa Basol, Jon Roozenbeek, Sander van der Linden: Good News about Bad News: Gamified Inoculation Boosts Confidence and Cognitive Immunity Against Fake News, Journal of Cognition, Volume.3, Issue.1, pp.1-9 (2020).