

フェイクニュースに対する警告タグがニュース認知に与える影響について

研究代表者

寶 雪

立命館大学 総合心理学部 准教授

1 研究目的

2016年のアメリカ大統領選以来、ソーシャルメディア（SNS）を媒介して拡散されるフェイクニュースへ注目が集まるようになった。最近では、新型コロナウイルスの流行に伴って、コロナ関連のフェイクニュースも見られており、これらのニュースから生じる恐怖感情は実際のウイルス以上に人々の行動に影響を与えることが指摘されている（Binjola & Patel, 2020）。うわさや偽情報が社会的に広がること自体は昔からよく見られる事象である。しかし、SNSの普及により、フェイクニュースはより遠く、速く、深く、広く拡散するようになり、問題を一層深刻化させている。今やSNSなどを通じて、根拠のない情報が大量に拡散されてしまう現象はインフォデミックと呼ばれ、その社会的な影響が懸念されている。それだけに、今後どのように対策を講じていくべきかについては即急に議論すべき課題だといえる。

こうした問題への対策の一つとして、近年では、海外を中心にフェイクニュースの見出しに警告タグをつける動きが出てきている。例えば、Google Chromeの拡張機能（Chromeに機能を追加するプログラムファイル）には、ユーザーが閲覧している情報が信頼できるか否かを判断し、その結果を提示するものが導入されている（Hinchliffe, 2016）。また、最近になってTwitter上においても誤解を与える情報を含んだツイートには何らかの警告ラベルが表示されるようになっており（BBC, 2020）、ウェブのプラットフォーム会社を中心に警告ラベルの活用が広がっている。

しかし、この警告タグの効果に関しては、実はまだわからない点も多い。例えば、いくつかの研究では警告タグをつけることで、フェイクニュースの信憑性評価を低められる「警告効果（Warning effect）」が見られているが（Chan, Jones, Jamieson, & Albarracín, 2017; Ecker, Lewandowsky, & Tang, 2010）、警告タグをつけることで、逆にフェイクニュースの信憑性が高まってしまう「バックファイア効果（Backfire effect）」も報告されている（Berinsky, 2017; Nyhan & Reifler, 2010）。加えて、サイト上に警告タグ付きニュースがあることによって、同じサイトにある警告タグがつかないニュースの信憑性評価にも影響があることもわかっている。これは閲覧者が警告タグの存在を認識している場合、逆に警告タグがついていないニュースについて、そのニュースはチェックされなかったためにタグがついていないのか、それとも正確なニュースだと検証されたためにタグがついていないのか、どちらであるかが曖昧であるがゆえに生じると考えられる（Pennycook et al., 2019）。結果として、警告タグがつかないことが安心材料となることでニュースに対する信憑性が高まったり、逆に警告タグの存在によってニュース全体に対して不信感を抱き、結果としてニュースの信憑性を低く見積もってしまったりすることがあり、これは警告タグの「波及効果（Spillover effect）」と呼ばれている（Pennycook et al., 2019）。これらのことから、警告タグがニュースの信憑性評価に及ぼす効果は、引き続き検討する必要がある。

加えて、SNSを介して拡散されることが多いフェイクニュースだが、SNS特有のメディア環境が警告タグの有効性に与える影響についてはこれまであまり議論されてこなかった。例として、SNSでの情報共有には「いいね!」、「リツイート」や「コメント」数が見える形で表示され、閲覧者はこれらを通してそのニュースに対する他人の反応を知ることができる。先行研究によると、こうした付随情報は私たちが情報の重要性や信憑性を判断する過程で一定の影響を及ぼすという。例えば他者の反応が多い場合には私たちはそのニュースは他人に気に入られたものであると考え、そのニュースが注目に値すると認識しやすくなる傾向がある（Avram, Micallef, Patil, & Menczer, 2020）。ここから、信憑性の低い情報であっても、他者の反応が多い場合にはその情報の信憑性を高く見積もる可能性があり、警告タグの警告効果があらわれにくいとも考えられる。

以上のことから、本研究では以下の三つの研究設問をたて検証する。

研究設問1：フェイクニュースに警告タグがつけられた場合、警告タグがついていない場合と比較して利用者のニュースに対する信憑性は低くなるだろうか（警告効果）、それとも逆に高くなるのだろうか（バックファイア効果）。

研究設問2：警告タグの存在は、タグがつかないその他のニュースの信憑性評価に影響を与えるだろうか（波

及効果)。

研究設問 3：研究設問 1 と 2 で見られるような警告タグの効果はニュースに付随する他者の反応数の違いによって、変化することはあるだろうか。

2 研究方法と結果

上記の研究設問を検証するため、三つの実験を行った。実験 1 は筆者が所属する大学の学生 192 名を対象にしたものであり、実験 2 と 3 は web 調査会社を介して集めた一般人 (実験 2:540 名、実験 3:654 名) を対象としている。3 つの実験デザインは基本的に同じであるが、実験 1 と 2 では実施後に改善点がいくつか見られたため、次に続く実験では若干の改良を行なっている。以下に三つの実験の詳細と結果、及び考察をまとめる。

2-1 実験 1

(1) 概要

2020 年 5 月から 7 月にかけて、大学生 192 名 (男子 43 名、女子 146 名、その他 3 名、年齢幅 18~30 歳、平均年齢 19.11) を対象に 2 (警告タグの存在有り・無し) x 2 (他者の反応数多い・少ない) の参加者間実験をオンライン上で行なった。参加者は研究参加に対して同意した後、ランダムに 4 つの群に分けられ、各群に割り当てられた教示文とニュース見出しを閲覧した。その後各々のニュースについて、信憑性の評価と以前に閲覧したことの有無について回答した。

(2) 刺激

今回実験刺激として 4 つのニュースを用いた。4 つともオンラインで公開されているものであるが、フェイクニュース刺激として用いた 2 つの見出し (F1・F2) は、既に信憑性がないと確認されたものであり、リアルニュース刺激として用いた 2 つの見出しは信頼できる情報源から発信されたものを用いた。各々の見出しについては表 1 にまとめた。なお、実験の中で、参加者に 4 つのニュースそれぞれについて、以前に見たことがあるかどうかを尋ねており、見たことがあると回答した人は除外した上で、各々のニュースを分析している。表 1 の右端の列には、見たことがあると回答した人を除外した参加者の人数、すなわち最終的に分析で使用した人数が記述されている。

表 1 実験 1 で使用したニュース見出し

| タイプ | 見出しの内容 | 警告タグ存在あり群でのタグ付けの有無 | 最終分析数 |
|--------------|-----------------------------------|--------------------|-------|
| フェイク ニュース | F1 「死亡事故を楽しむ日本人」と CNN が報道 | 有り | 167 |
| | F2 NHK の人件費は年々増額 | 無し | 167 |
| リアル ニュース | R1 パワハラで自殺教唆疑い 兵庫県警、三菱電機 30 代社員を… | 無し | 148 |
| | R2 日米新貿易協定受け政策大綱を決定 和牛の増産支援など | 無し | 157 |

(3) 警告タグ条件の操作

警告タグの存在有り群：まず教示文のところで、警告タグの存在を知らせる文章 (「第三者機関によるチェックの結果、内容の信憑性に問題があると判断されたニュースには、下記のような警告タグがついています。)」と実際の警告タグの写真 (図 1) を見てもらった。次に警告タグがついたフェイクニュース 1 件 (F1、図 2)、警告タグがつかないフェイクニュース 1 件 (F2)、及びリアルニュース 2 件 (R1、R2) を閲覧してもらった。なお、F1 は常に最初に表示され、残り 3 件のニュースはランダムに表示されるように設定した。

警告タグの存在無し群：警告タグの存在をふせたまま、警告タグをつけていない状態のフェイクニュース 2 件 (F1 と F2) とリアルニュース 2 件 (R1、R2) を閲覧してもらった。なお、ニュースの提示順については警告タグあり群と同様に、F1 を最初に提示し、残り 3 件をランダムに提示した。



第三者機関によるチェックの結果、内容の信憑性に問題あり
もっと見る

図1 3つの実験で使用した警告タグ



図2 警告タグの存在有り・他者の反応が多い群のフェイクニュースの例（実験1, F1）

（4）他者の反応数条件の操作

他者の反応数が少ない群のコメント数を0～15、リツイート数を0～50、いいね数を0～100に設定した。それに対し、他者の反応数が多い群のコメント数を100～150、リツイート数を1500～3000、いいね数を3000～5000に設定した。

（5）尺度

ニュースの信憑性を測る尺度としてAppelman and Sundar (2015) のメッセージ信憑性尺度を用いた。具体的には、「信用できる」、「本物である」、「正確である」という三つの設問を五件法（1.当てはまらない～5.とても当てはまる）で回答してもらった。各ニュースそれぞれのクロンバックの α 係数を算出した結果、F1が.89、F2が.89、R1が.90、R2が.90であった。

（6）実験1の分析結果

SNSでの他者の反応が警告タグの有効性に与える影響を検討するため、警告タグの存在有り・無し条件と他者の反応数多い・少ない条件を独立変数、ニュースの信憑性を従属変数、性別、年齢を統制変数として、共分散分析を行った。表2にあるように、F1において、警告タグの存在有り群（ $M = 1.86$ ）の方が、存在無し群（ $M = 2.21$ ）よりもニュースに対する信憑性が低くなった。加えてR1についても、警告タグ存在有り群（ $M = 3.81$ ）の方が、存在無し群（ $M = 3.54$ ）よりも10%の水準で信憑性が有意に高かった。一方で、他者の反応の主効果、交互作用は見られなかった。

（7）実験1の考察

F1の結果より、フェイクニュースに警告タグを付けることで、ニュースの信憑性評価が低くなったことから、警告効果が見られたといえる。また、R1の結果から、警告タグの存在を知っていることで、リアルニュースの信憑性を高く見積もってしまう傾向が見られた。これは、波及効果の一種であるといえる。一方で、今回は他者の反応数の違いによって、ニュース見出しの信憑性評価に違いが現れなかった。一つ考えられる理由として、反応数の設定が不十分だった可能性がある。実験中に行われた操作チェックの結果では、反応

数多い群に振り分けられた参加者のうち、平均して約 42%の人が、閲覧したニュースについている他者の反応数が少ないと回答していた。ここから、反応数が多い群の操作においては、いいね数、リツイート数、コメント数をもっと多く設定する必要があることが示唆される。また、今回の参加者は大学生に限定されており、男女比にも大きなばらつきがあった。そのため、実験 2 では反応数の値を調整し、さらに参加者の年齢幅と男女比も調整する必要があると結論づけた。

表 2 実験 1 の共分散分析結果のまとめ

| | 警告タグ条件 | | 他者の反応条件 | | | | 交互作用 | |
|----|--------|------|---------|------|-------|-----|--------|--|
| | 存在あり群 | | 存在なし群 | | 反応多い群 | | 反応少ない群 | |
| | M | M | F | M | M | F | F | |
| F1 | 1.86 | 2.21 | 7.63 ** | 2.04 | 2.03 | .04 | 1.99 | |
| F2 | 2.81 | 2.65 | 1.78 | 2.71 | 2.76 | .21 | .58 | |
| R1 | 3.81 | 3.54 | 3.71 † | 3.67 | 3.67 | .02 | .81 | |
| R2 | 3.94 | 3.83 | .53 | 3.86 | 3.91 | .17 | .73 | |

** $p < .01$, † $p < .10$

2-2 実験 2

(1) 概要

2020 年 11 月にウェブ調査会社を介して集めた 540 名に対して 2 (警告タグの存在有り・無し) x 2 (他者の反応数多い・少ない) の参加者間実験をオンライン上で行なった。その際、参加者の年齢、性別については 20 代～50 代の男女それぞれを 64 名、または 63 名になるように割り付け設定を行った。データ収集後、回答に不備のある 60 名を除外し最終的に 476 名 (男性 237 名、女性 239 名) を分析対象とした。参加者の年齢は 20 歳から 60 歳までであり、平均年齢は 39.74 歳 ($SD = 11.54$) であった。実験 1 と同様、参加者は 4 つの群にランダムに分けられ後、各群に割り当てられた教示文とニュース見出しを閲覧し、その後各々のニュースについて、信憑性の評価と以前に閲覧したことの有無について回答した。

(2) 刺激

実験 1 ではフェイクニュース 2 つ、リアルニュース 2 つの計 4 つのニュース見出し刺激を使ったが、実験 2 ではフェイクニュース 4 つ、リアルニュース 2 つの計 6 つの刺激を使用した。実験 1 と同様、6 つ全てがオンラインで公開されているものであり、フェイクニュース刺激として用いた 4 つの見出しは、既に信憑性がないと確認されたものであり、リアルニュース刺激として用いた 2 つの見出しは信頼できる情報源から発信されたものを用いた。表 3 に具体的に利用した刺激の見出しをまとめた。

表 3 実験 2 と実験 3 で使用したニュース見出し

| タイプ | 見出しの内容 | 警告タグ存在あり群でのタグ付けの有無 | 最終分析数 | |
|--------------|---------------------------|--------------------|-------|------|
| | | | 実験 2 | 実験 3 |
| フェイク ニュース | F1 感染対策に緑茶が有効 | 有り | 441 | 547 |
| | F2 習近平氏の娘がハーバードの卒業式で歌う | 有り | 459 | 582 |
| | F3 WHO が方向転換「感染者の隔離は不要」 | 無し | 443 | 571 |
| | F4 安倍首相、持病再発後も「焼肉食いまくってた」 | 無し | 430 | 557 |
| リアルニ ュース | R1 コロナ「陰性」証明、ニーズ高まる | 無し | 407 | 542 |
| | R2 デジタル化への集中投資に 1158 億円 | 無し | 397 | 515 |

(3) 警告タグ条件の操作

警告タグの存在有り群：実験 1 と同様、この群の参加者には最初の教示のところで警告タグの存在を知らせる文章と実際の警告タグの写真を提示した (図 1)。その後、警告タグがついたフェイクニュース 2 件のうち 1 つをランダムに提示した (F1 または F2)。続いて、先ほど提示されなかった方の警告タグ付きフェイクニュース 1 件 (F1 または F2) と警告タグがつかないフェイクニュース 2 件 (F3 と F4)、及びリアルニュース 2 件 (R1 と R2) をランダム順で全て提示した。

警告タグの存在無し群：警告タグの存在を伏せたまま、ニュース見出し 6 件を提示した。また F1 と F2 には警告タグを付けずに参加者に提示した。なお、ニュース見出しを提示する順番は警告タグ有り群と同様の

手順を用いた。

(4) 他者の反応数条件の操作

実験1での操作チェックの結果を踏まえ、反応が多い群のコメント数を1000~4000、リツイート数を1500~6000、いいね数を40000~100000に設定した。なお、反応が少ない群に関しては、実験1と同様にコメント数を0~15、リツイート数を0~50、いいね数を0~100に設定した。

(5) 尺度

実験1と同様にAppelman and Sundar (2015) のメッセージ信憑性尺度を用いた。各ニュースそれぞれのクロンバックの α 係数を算出した結果、いずれも.94から.96の範囲内に収まっており、尺度の信頼性は問題ないと判断した。

(6) 実験2の分析考察

実験1と同様、警告タグの存在有り・無条件と他者の反応数多い・少ない条件を独立変数、各ニュースの信憑性を従属変数、性別、年齢を統制変数として、共分散分析を行った(表4)。結果として、警告タグの主効果、他者の反応数の主効果、交互作用全てにおいて、有意な結果が得られなかった。一方で、年齢がニュース見出しの信憑性評価に有意な影響を与えていた。具体的にはF3において、年齢が高いほどニュース見出しの信憑性を低く評価していた($F = 14.47, p < .001$)。

表4 実験2の共分散分析結果のまとめ

| | 警告タグ条件 | | | 他者の反応条件 | | 交互作用 | |
|----|--------|-------|------|---------|--------|------|------|
| | タグあり群 | タグなし群 | F | 反応多い群 | 反応少ない群 | F | F |
| | M | M | | M | M | | |
| F1 | 1.95 | 2.00 | .37 | 1.97 | 1.97 | .00 | .01 |
| F2 | 2.26 | 2.36 | .97 | 2.28 | 2.33 | .44 | .05 |
| F3 | 2.17 | 2.09 | 1.00 | 2.10 | 2.15 | .44 | 2.72 |
| F4 | 2.21 | 2.11 | 1.33 | 2.15 | 2.18 | .16 | .09 |
| R1 | 2.83 | 2.75 | .80 | 2.78 | 2.78 | .02 | .05 |
| R2 | 2.97 | 2.84 | 2.01 | 2.95 | 2.86 | .18 | 2.35 |

** $p < 0.01$, * $p < 0.05$, † $p < 0.10$

(7) 実験2の考察

実験1で見られた警告効果、及び波及効果が今回の実験2では見られなかった。考えられる理由として、実験2では参加者を20歳~60歳の一般人に広げたため、今回の実験結果に影響を与えるうる変数が紛れ込んでしまい、結果を見えづらくした可能性がある。例として、SNS利用頻度に関する変数が考えられる。本実験では、他者の反応数というSNSに関連した独立変数を含んでいるが、情報の信憑性判断をするに際して、他者の反応数を参考にするという習慣はSNSをある程度利用しなければ身につかない可能性がある。また、フェイクニュースはSNS上で広がる傾向が多いため、普段からSNSを介してニュースに接触している人は、そうでない人に比べ、フェイクニュースに遭遇している確率が多く、その分警告タグに対しても敏感に反応する可能性がある。SNSでニュースを閲覧する頻度は大学生のみを参加対象とした実験1ではばらつきが少なく、あえて統制変数として分析に含めなくとも結果に影響を与えなかったと考えられる。一方で、一般人に参加対象を広げた際には、参加者間のSNS利用の違いによって、警告タグと他者の反応の効果が見えにくくなった可能性も否めない。そこで、続く実験3では参加者が日頃どのくらいSNS上でニュースを閲覧しているかを測定し、統制変数として分析に入れる。

2-3 実験3

(1) 概要

2021年2月にウェブ調査会社を介して集めた654名に対して実験2と同様の2(警告タグの存在有り・無し)x2(他者の反応数多い・少ない)の参加者間実験をオンライン上で行なった。参加者の年齢、性別については20代~60代の男女それぞれを74名、または75名になるように割り付け設定を行った。データ収集後、回答に不備のある36名を除外し最終的に618名(男性307名、女性311名)を分析対象とした。参加者の年齢は20歳から60歳までであり、平均年齢は39.46歳($SD = 11.49$)であった。

実験の手続き、刺激として用いたニュース見出し、警告タグ条件の操作、他者反応数条件の操作については実験2と同様のものを用いた。

(2) 尺度

実験2と同様 Appelman and Sundar (2015) のメッセージ信憑性尺度を用いた。各ニュースそれぞれのクロンバックの α 係数を算出した結果、いずれも.94から.95の範囲内に収まっており、尺度の信頼性は問題ないと判断した。また、SNSでのニュース利用頻度を測るために、実験の終盤で日頃国内外のニュースを入手する手段としてSNSをどのくらい利用しているかについて参加者に尋ね、1(あまり利用しない)~5(とてもよく利用する)の中から当てはまるものを一つ選択してもらった。

(3) 実験3の分析結果

警告タグの存在有り・無し条件と他者の反応数多い・少ない条件を独立変数、ニュースの信憑性を従属変数、性別、年齢、SNSニュース利用頻度を統制変数として、共分散分析を行った(表5)。その結果、警告タグの主効果がF1($F = 4.30, p < .05$)とF2($F = 7.06, p < .01$)のニュース見出しで見られた。具体的には、どちらのニュース見出しにおいても、警告タグの存在有り群の方が存在無し群よりもニュースの信憑性を低く見積もっていた。加えて、他者の反応数の主効果についてもF1($F = 4.48, p < .05$)では5%水準で、F2($F = 2.79, p < .10$)では10%水準で有意な結果が見られた。どちらのニュース見出しにおいても、他者の反応数が多い群の方が少ない群に比べニュースの信憑性を高く見積もっていた。

加えて、警告タグと反応数の交互作用もF3とR2のニュース見出しにおいて10%水準で有意であった。F3について詳しく見てみると、ニュース見出しの他者反応数が少ない場合、警告タグの存在有り群参加者は存在無し群の参加者に比べ、ニュースの信憑性をある程度高く見積もっていた。しかし、他者反応数が多い場合、タグ存在あり群の信憑性評価が下がり、逆にタグ存在無し群の信憑性評価は上がったため、結果としてタグ条件間の差はなくなっていた(図3)。R1について見てみると、反応数が少ない場合にはF3と同じく警告タグの存在あり群の参加者の方が存在なし群の参加者に比べニュース信憑性を高く見積もっていたが、他者反応数が多い場合には逆転しており、タグ無し群の方がタグあり群よりもニュース信憑性の得点が高かった(図4)。

また、統制変数として分析に組み入れていたSNSのニュース利用頻度を見ると、6つ全てのニュース見出しにおいて1%水準で有意な結果が得られた。具体的には、SNSのニュース利用頻度が高い人ほど、ニュース信憑性を高く見積もっていた(β 値の結果は、F1が0.15、F2が0.29、F3が0.18、F4が0.20、R1が0.21、R2が0.18であった)。

表5 実験3の共分散分析結果のまとめ

| | 警告タグ条件 | | F | 他者の反応条件 | | F | 交互作用 | |
|----|--------|-------|---------|---------|--------|--------|------|---|
| | タグあり群 | タグなし群 | | 反応多い群 | 反応少ない群 | | F | F |
| | M | M | | M | M | | | |
| F1 | 2.18 | 2.02 | 4.30 * | 2.02 | 2.18 | 4.48 * | .76 | |
| F2 | 2.39 | 2.20 | 7.06 ** | 2.23 | 2.35 | 2.79 † | .00 | |
| F3 | 2.01 | 2.12 | 2.09 | 2.06 | 2.07 | .02 | 2.73 | † |
| F4 | 2.18 | 2.25 | .82 | 2.18 | 2.26 | 1.12 | 1.81 | |
| R1 | 2.71 | 2.77 | .48 | 2.72 | 2.76 | .35 | 1.90 | |
| R2 | 2.71 | 2.71 | .00 | 2.69 | 2.74 | .37 | 3.47 | † |

** $p < 0.01$, * $p < 0.05$, † $p < 0.10$

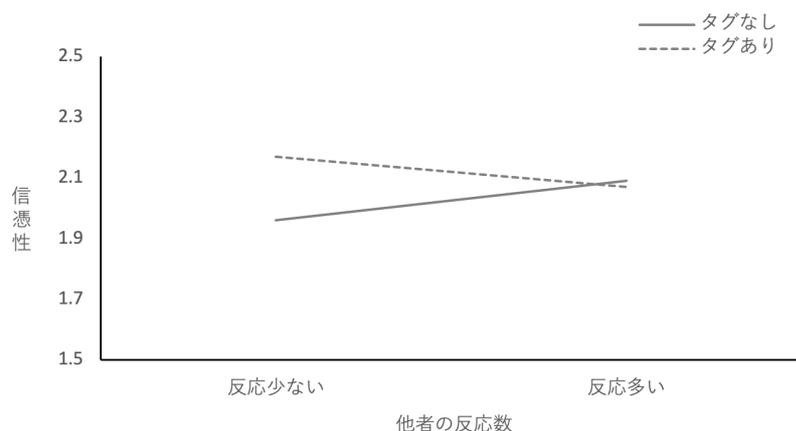


図3 F3における警告タグと他者の反応数の交互作用

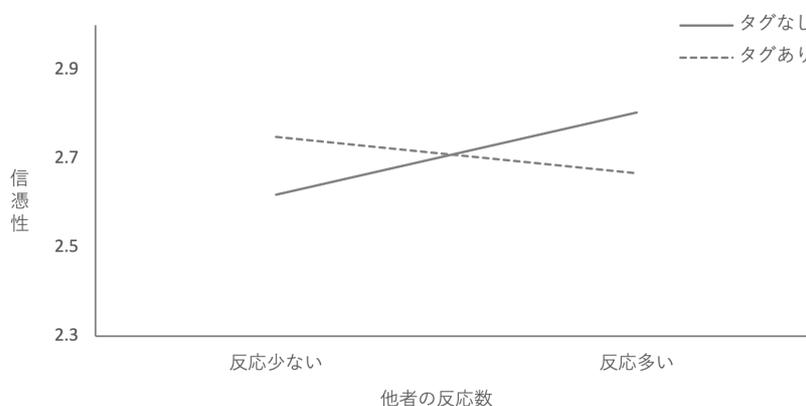


図4 R2における警告タグと他者の反応数の交互作用

(4) 実験3の考察

今回フェイクニュースの刺激としてF1～F4の4つのニュース見出しを参加者に提示しているが、F1とF2に関しては警告タグの警告効果を検証するために、警告タグの存在有り群にのみ図1のようなタグをつけている。結果として、警告タグの主効果がF1とF2の両方に見られたことから、タグをつけることで、フェイクニュースに対する信憑性を低める、いわゆる警告効果が確認されたといえる。一方で、この警告効果が、他者の反応を考慮した際にどのように変化するかをみると、交互作用がF1とF2ともに有意ではなかったことから、そのニュースに対する他者の反応数はタグの警告効果の有効性に影響を与えないことが示された。

次に、波及効果について考察すると、F3とF4、及びR1とR2においてタグ条件の主効果が見られていないことから、一見波及効果は確認できなかったように見える。しかし、この図式に他者の反応数という変数を入れると結果は若干異なってくる。具体的に、F3とR2の交互作用の結果より、他者の反応数が小さい状況下においては、タグの存在を知っていることで、タグがつかないニュースの信憑性を高く見積もってしまう波及効果が現れる可能性はあるといえる。ただ、他者の反応数が多くなると、そうした波及効果は消えてしまうか(F3の場合)、もしくはタグの存在がタグのつかないニュースの信憑性を低めてしまう、逆の波及効果がみられた(R2の場合)。この理由については、次の総合考察で詳しく説明する。

3 総合考察

今回、警告タグの効果、及びこの警告タグ効果に他者の反応数が与える影響を明らかにするため、大学生を対象とした実験(実験1)と一般人を対象とした2つの実験(実験2、3)を行った。結果として、警告タ

グをつけることでフェイクニュースの信憑性を低下させる警告効果が実験1と実験3で見られた。また、警告タグの波及効果についても、実験1と実験3の両方で見られており、警告タグの存在を知っている状態において、タグがついていない他のニュース(そのニュースがフェイクであるかリアルであるかにかかわらず)の信憑性が変化する傾向があることが確認された。

本研究では三つの研究設問をたてているが、研究設問1と2は警告タグがニュースの信憑性評価に与える影響(警告効果、バックファイア効果、波及効果)についての検討であり、研究設問3は他人の反応数が警告タグの影響をどのように変化させるかについてであった。まず、研究設問1の検討課題であるフェイクニュースに警告タグをつけることの有効性について考察すると、今回は警告タグがついたニュース見出し(実験1のF1、実験3のF1・F2)において警告効果がみられ、逆に警告効果と表裏の関係にあるバックファイア効果は確認されなかった。よって、本研究の結果を見る限り、警告タグはフェイクニュースの信憑性を正しく認知する上で、ある程度の役割を果たすといえる。ただ、今回、警告効果のみがみられた理由として、本研究の刺激として用いたニュースの性質が関係している可能性もある。先行研究においてバックファイア効果が見られたものの多くは、刺激として政治に関するニュースを用いており、また特に米国で実施されたものが多かった。米国では政治に対する考え方がリベラルと保守という2つの立場に大きく分かれており、自分の政治思想に沿ったニュースに警告タグがついた場合は、タグに反発を感じ、逆にそのニュースを支持する傾向が強まることがいくつかの研究で報告されている(Nyhan & Reifler, 2010)。ここから、バックファイア効果が出出するためには、そのニュースの内容について、参加者が予めある程度の支持を形成している必要がある。今回は参加者がタグ付きフェイクニュース(実験1のF1、実験2と3のF1・F2)の内容について、中立的な立場であったと考えられ、その結果バックファイア効果が出出しなかったという可能性も否めない。今後は参加者のニュースに対する事前の態度を操作、もしくは測定した上で、バックファイア効果の検討をする必要があるといえる。

また、今回の実験では、SNS上での「いいね!」、「リツイート」、「コメント」のような他者の反応が大きい場合、警告効果は小さくなると予想していたが、全ての実験において交互作用が見られなかった。このような結果になった理由として、二つの可能性が考えられる。一つ目は、そもそも他者の反応がニュース信憑性の判断に影響を及ぼさない可能性である。ただ、これに関しては、実験3において他者の反応数が多い方群にいた参加者の方が、少ない群にいた参加者よりもF1とF2の信憑性を高く評価していたことから、理由としては考えられにくい。二つ目は、警告タグと他者の反応の両方が並んだ場合、警告タグの方が信憑性を判断する手がかりとしてより好まれたという可能性である。すなわち、ニュース見出しを見た時点において、参加者は警告タグと他者の反応数の二つの手がかりを使ってニュース信憑性を評価したが、警告タグの方がより信頼できる判断材料であると認知し、警告タグを主な手がかりとしてニュース信憑性を評価したという心理的なプロセスが想定される。ここから、警告タグへの信頼度が高い状況においては、他者の反応が高くとも、警告効果は保たれることが示唆される。

次に、研究設問3であげた警告タグの波及効果について考察してみたい。今回実験中にタグを付けないフェイクニュース(実験1のF2、実験2と3のF3・F4)とリアルニュース(R1とR2)を設定し、タグの存在を知っていることで、タグがつかないその他のニュースの信憑性が変化する波及効果について検証した。結果として、実験1と実験3において、一部のニュース見出しに波及効果が確認されたが、興味深いことに、実験3の結果から他者の反応数によって波及効果の表れ方に違いが生じたことがわかった。具体的には、他者反応数が少ない条件では、タグがつかないフェイクニュース(F3)とリアルニュース(R2)の信憑性が低くなっていたが、逆に他者反応数が多い条件においては、その効果が消えていたり(F3の場合)、R2にいたっては、タグの存在を知っている群の方がR2の信憑性を低く見積もるといって、先ほどとは真逆の波及効果が生じていた。あくまでも予測ではあるが、参加者は次のような心理プロセスを辿ったと考えられる。つまり、反応数が少ない場合、タグの存在を知っていた参加者は、ニュース見出しに警告タグがないという事実を主要な手がかりに「警告タグがない=フェイクニュースでない」という判断を下し、結果としてニュースの信憑性を高めに評価したのではないだろうか。逆に反応数が多い場合、ニュース見出しに警告タグがないという事実に加え、他者の反応数も判断材料として加わった結果、反応数が少ない条件とは異なった波及効果が表出したと考えられる。もっとも、この状況において、他者反応数が多いという事実は、当初予想していたような「ニュースの他者反応数が多い=他者に注目されているニュースである=このニュースは信頼できる」というような図式で効果を発揮したのではなく、他にもフェイクニュースがあるかもしれないという懷疑心から「ニュースの他者反応数が多い=他者に広く拡散されたニュースである=フェイクニュースの可能性も

高い」という図式に沿って、参加者心理に働いたと推測される。その結果、タグの存在を知っていた群においては、他者反応数が多い時の方が、参加者のニュースに対する信憑性が低下したと考えられる。一方で、タグの存在を知らされていない群の参加者は、他者の反応数を主な手がかりとしてニュースの信憑性判断を行い、結果として「ニュースの他者反応数が多い＝他者に注目されているニュースである＝このニュースは信頼できる」という図式にそった形でニュースの信憑性評価をしたと考えられる。このことから、他者の反応数が多いという情報は、状況によってはフェイクニュースの特徴の一つとして参加者に認識されることが示唆される。

上記の結果から、ニュースの信憑性評価において、参加者は様々な手がかりをたよりに、複合的に判断を下していることがうかがえる。そのため、警告タグの効果はそれ単体で評価するのではなく、同じ文脈で提示されるその他の情報の影響も視野に入れながら、総合的に評価する必要があるといえる。本研究の結果から、閲覧者が警告タグそのものに信頼をおいている場合であれば、フェイクニュースの信憑性を正す上で、警告タグはある程度の効力をもちえることがわかった。このことから、警告タグの利用は促進されるべきだといえる。ただ、他者の反応数もニュースの信憑性判断において一定の影響をもつことも念頭に入れる必要がある。そのため、もし間違っ警告タグをつける事例などが発生し、警告タグに対する信頼性が低くなってしまった状況では、他者の反応数の影響が強くなってしまいう可能性も否めない。その結果、反応数が多いほどニュース信憑性が高まってしまい、拡散されたフェイクニュースを正すことはもっと難しくなってしまうとも考えられる。加えて、警告タグを多用することで、タグがつかないその他のニュースの信憑性を歪める可能性があることも考慮する必要がある。ここから、警告タグは闇雲に多用せず、期間限定（政治的なフェイクニュースが流れやすい選挙期間中だけ）、もしくはトピックに限定して（例えば、健康に著しい影響を及ぼす可能性のあるフェイクニュースに限定する）使用することを検討するのが良いといえる。今後は引き続きバックファイア効果、及びニュースのトピックや内容によって警告タグの有効性がどのような変化するかなどについて検討し、インフォデミックという社会問題に対して、有効な手立てを講じていくことが重要である。

【参考文献】

- Appelman, A., & Sundar, S. S. (2016). Measuring message credibility: Construction and validation of an exclusive scale. *Journalism & Mass Communication Quarterly*, *93*, 59-79.
- Avram, M., Micallef, N., Patil, S., & Menczer, F. (2020). Exposure to Social Engagement Metrics Increases Vulnerability to Misinformation. *HKS Misinformation Review* *1*, 1-9.
- BBC. (2016). 【英国民投票】離脱派が勝った 8 つの理由. Retrieved from: <https://www.bbc.com/japanese/features-and-analysis-36628343>
- Berinsky, A. J. (2017). Rumors and health care reform: Experiments in political misinformation. *British Journal of Political Science*, *47*, 241-262.
- Binjola, H., & Patel, K. (2020). Fake News Swamping Interpersonal Communication in the Times of Corona Virus. *International Journal on Transformations of Media, Journalism & Mass Communication*, *5*, 1-6.
- Chan, M. P. S., Jones, C. R., Hall Jamieson, K., & Albarracín, D. (2017). Debunking: A meta-analysis of the psychological efficacy of messages countering misinformation. *Psychological science*, *28*, 1531-1546.
- Ecker, U. K., Lewandowsky, S., & Tang, D. T. (2010). Explicit warnings reduce but do not eliminate the continued influence of misinformation. *Memory & cognition*, *38*, 1087-1100.
- Hinchliffe, E. (2016). This Chrome extension will warn you when you visit a fake news site. Retrieved from: <https://mashable.com/article/fake-news-alert-chrome-extension>
- Nyhan, B., & Reifler, J. (2010). When corrections fail: The persistence of political misperceptions. *Political Behavior*, *32*, 303-330.
- Pennycook, G., Bear, A., Collins, E., & Rand, D. G. (2019). The Implied Truth Effect: Attaching Warnings to a Subset of Fake News Headlines Increases Perceived Accuracy of Headlines Without Warnings. *Management Science*, 4944-4957.

Pennycook, G., & Rand, D. G. (2019). Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition*, 188, 39-50.

〈発 表 資 料〉

| 題 名 | 掲載誌・学会名等 | 発表年月 |
|--|---------------------------|-------------|
| フェイクニュースに対する警告タグは有効なのか ―日本大学生を対象とした検証― | 2020 年度春季（第 42 回）情報通信学会大会 | 2020 年 7 月 |
| 情報通信機器への愛着がフェイクニュース警告タグに与える影響 | 日本社会心理学会第 61 回大会 | 2020 年 11 月 |
| SNS 上における他者の反応が警告タグの有効性に与える影響 | 2020 年度秋季（第 43 回）情報通信学会大会 | 2020 年 12 月 |
| | | |