

確証バイアスとウェブ検索行動の関係分析

鈴木 雅貴[†] 山本 祐輔[†]

[†] 静岡大学大学院総合科学技術研究科 〒432-8011 静岡県浜松市中区城北 3-5-1

E-mail: [†]suzuki@design.inf.shizuoka.ac.jp, ^{††}yamamoto@inf.shizuoka.ac.jp

あらまし 本研究では、自身の意見や信念を支持する情報を優先的に閲覧してしまう確証バイアスとウェブ検索行動の関係について分析を行った。具体的には、オンラインユーザ実験にて、実験協力者に対して検索トピックの印象を操作する異なった事前情報を提示することで、擬似的に確証バイアスを発生させ、検索タスク中の行動ログの分析を行った。ユーザ実験の結果、検索トピックに対して否定的な信念を持つウェブ検索ユーザは、批判的に情報を精査するスキルが低い場合、確証バイアスによる負の影響を受けやすいため、検索結果一覧ページを閲覧する時間が短く、上位の検索結果をクリックし、肯定的な意見を閲覧しない傾向にあった。本研究で得られた知見によって、確証バイアスを考慮したウェブ情報の精査行動を促進するシステムに寄与することが期待される。

キーワード 情報検索、情報信頼性、確証バイアス、ヒューマンファクター

1 はじめに

近年、ウェブ情報の信憑性が問題になっている。たとえば、健康情報に関するウェブ上の情報は、その半分以上が専門家のチェックを受けていないことが明らかにされている [1]。それゆえ、ユーザはウェブ情報が正しいか否かを見分けられなければ、誤った情報を鵜呑みにしてしまう可能性がある。

ウェブ情報の信憑性に関する問題は、ウェブ検索エンジンやレコメンドシステムといった情報配信の個人最適化技術によって助長されている。情報配信の個人最適化によって、ユーザが興味関心のある情報にしかアクセスしなくなる現象「フィルターバブル」は、慎重な意思決定を行うために広い視野で情報を精査する機会を奪うため、社会問題になりつつある [2]。

誤った情報、品質が低い情報を信じてしまう背景には、認知心理学の分野で知られている「確証バイアス」が関係している。確証バイアスとは、「無意識のうちに自分の意見や仮説を支持するような情報を優先的に探す」という人間の認知傾向である [3]。例えば、健康意識の高いユーザ X が、テレビ放送で遺伝子組み換え技術が使用された食品 Y は健康に被害があるといったことを耳にし、食品 Y に不信感を抱いたとする。ユーザ X が食品 Y の安全性についてウェブ検索を行うと、ユーザ X は「食品 Y は健康に被害がある」と考えたいがために、たとえ信憑性が低かったとしても食品 Y に否定的な情報を優先的に閲覧してしまう。

このように、確証バイアスは、ウェブ検索行動においても大きな問題となりえる。特に、衣食住や政治に関してウェブ検索する際に発生する確証バイアスは、社会生活に大きな影響を与えかねない。実際に、健康トピックのウェブ検索時に確証バイアスが発生しうることが、White によって明らかにされている [4]。それゆえ、ユーザが信憑性が低い情報を取得しないようにするためには、確証バイアスの存在を考慮した情報アクセスシステムの設計が重要となる。本研究では、ウェブ情報の精

査行動を促進するシステムを設計するために、確証バイアスとウェブ検索行動の関係を定量的に明らかにする。

情報検索分野や HCI 分野では、情報探索時の確証バイアスを緩和するための研究がなされている [5] [6]。その一方、確証バイアスを有するウェブ検索ユーザがどういった検索行動を取るか明らかにする研究はあまり行われていない。確証バイアスを有するウェブ検索ユーザが行う典型的な行動が明らかになれば、それを考慮した情報アクセスシステムの設計が可能になる。

本研究では、実験協力者に与える検索トピックに関する事前情報の印象を操作することによって、確証バイアスを擬似的に発生させ、確証バイアスを有するユーザとそうではないユーザとのウェブ検索行動の違いを分析した。本研究では、以下の仮説を設定し、オンラインユーザ実験を行った：

H1 確証バイアスを有するウェブ検索ユーザは、どの情報を詳細に見るべきかに注意を払わないため、検索結果一覧を閲覧する時間が短く、検索結果上位のページを優先的に閲覧する。

H2 確証バイアスを有するウェブ検索ユーザは、情報を注意深く閲覧しようとしなため、ウェブページの閲覧時間が短く、ウェブページの閲覧数が少ない。

H3 仮説 H1 および H2 の程度は、ウェブ検索ユーザの批判的思考能力の多寡にも影響を受ける。

2 関連研究

2.1 ウェブ検索における確証バイアス

認知心理学では、自身の意見や仮説に一致する情報を優先的に閲覧する「確証バイアス」が、意思決定に多大な影響を及ぼすことが知られている [7]。ウェブ検索においても、確証バイアスの存在とその影響が明らかにされている。

White はウェブ検索において、ユーザの信念がウェブ検索に与える影響について分析を行った [4]。その結果、ウェブ検索ユーザのもつ事前信念が弱い場合、ウェブ検索を経ることで事

前信念が強化されやすいことを明らかにした。また、ウェブ検索ユーザはポジティブな検索結果に影響を受けやすいことを明らかにした。Pothirattanachaikul らは文書の言明極性と信憑性が、ウェブ検索ユーザの検索行動と信念の変化に及ぼす影響について研究を行った。その結果、自分の信念と一貫しない意見の書かれたウェブページを見たユーザは、より検索行動に時間をかけるということを明らかにした [8]。Liao らは、医薬品に対するコメントが「副作用」または「効果」について、どの程度述べているか可視化したシステム ASPECT INDICATORS を提案した。さらに、ユーザ実験の結果、医薬品の検索タスクにおいて、提案システムが検索ユーザの確証バイアスを軽減することを示した [6]。

これらの研究から、ウェブ検索において確証バイアスが影響を与えるといえる。しかし、確証バイアスを有するウェブ検索ユーザの検索行動を分析した研究は少ない。そのため、本研究では実験協力者に対して擬似的に確証バイアスを発生させることで、確証バイアスとウェブ検索行動の関係を定量的に明らかにする。

2.2 ウェブ検索ユーザの検索・閲覧行動の分析

行動ログを用いてウェブ検索行動の分析を行う研究は数多くなされている。White らは、検索演算子を用いてウェブ検索ユーザの検索熟練度を定義し、検索熟練度と検索行動の関係性に関するログ分析を行った [9]。分析の結果、検索熟練度によって検索結果クリック数やセッション時間に差が生じることを明らかにした。Carla らは、ヘルスリテラシーとウェブ検索行動の関係性について目の動きを用いて分析した [10]。その結果、ヘルスリテラシーの高いウェブ検索ユーザは、より多くのウェブサイトを訪問し、文章の閲覧に多くの時間を費やすことが明らかになった。

2.3 批判的思考・態度

Kusumi らによると、情報を適切に判断するためには、言語能力や推論力といった批判的思考スキルだけでなく、批判的であろうとする批判的思考態度が必要とされている [11]。また、Petty らが提唱した精査可能性モデルによると、情報内容を精査するスキルを活かす前提として、情報精査への動機付けが必要とされている [12]。確証バイアスは情報精査態度に負の影響を与えると考えられる。以上のことから、ウェブ検索ユーザの情報精査を促進するには、ユーザの確証バイアスを緩和、軽減することが重要となる。そのため、本研究では、確証バイアスに加えて、批判的思考能力がユーザのウェブ検索・閲覧行動に及ぼす影響について調べた。

3 実験方法

本章では、確証バイアスがウェブ検索行動に与える影響を分析する方法について述べる。本研究では、オンライン上で健康トピックに関するウェブ検索タスク実験を行った。本実験では、実験協力者に検索トピックに対する事前信念を強める情報を与えることで、擬似的に確証バイアスを発生させ、その後の検索

行動への影響を調査した。

以下、実験の詳細について記す。なお、本研究では、検索トピックに対してネガティブな信念を有する群を **biased(-)** 群、ポジティブな信念を有する群を **biased(+)** 群、どちらでもない群を **critical** 群と呼称する。

3.1 実験協力者

クラウドソーシングサービス Lancers.jp¹を用いて、300 名の実験協力者を募った。タスクが完全に終了していない者、何らかの理由で複数回タスクに取り組んでいる者、ウェブページを閲覧していない者のデータは除外した。最終的に、合計 275 名の実験協力者のデータを分析に用いた。

各実験協力者は、検索タスク開始前に検索トピックに対する事前信念の強さを回答した。実験協力者は、「1. 危険である」「2. やや危険である」「3. どちらでもない」「4. やや安全である」「5. 安全である」の選択肢から回答を行った。その後、事前信念に応じて実験協力者を特定の群に割り当てた。実験協力者の内訳は、**biased(-)** 群が 148 人、**critical** 群が 96 人、**biased(+)** 群が 31 人である。タスクを完了した実験協力者には、報酬として 100 円を支払った。

3.2 検索トピック

本研究では、検索トピックに関する事前信念が重要になる。そのため、検索トピックとして、賛否両論のある「遺伝子組み換え食品」を選んだ。各実験協力者は、「遺伝子組み換え食品を摂取することの安全性」についてウェブ検索で調査するタスクに取り組んだ。

3.3 実験手順

まず、実験協力者は Lancers.jp にてユーザ登録したあと、研究室で用意した実験用サイトに移動した。

次に、実験協力者は検索トピックに関する事前知識、事前信念を問うアンケートに回答した。事前知識を問うアンケートでは、「遺伝子組み換え食品についてどの程度知っていますか？」という質問を課した。実験協力者は、「1. まったく知識がない」「2. あまり知識がない」「3. そこそこ知識がある」「4. かなり知識がある」「5. 十分に知識がある」の選択肢から回答を行った。また、事前信念を問うアンケートとして、実験協力者には「遺伝子組み換え食品を摂取することの安全性についてどのように考えますか？」という質問に 5 段階のリッカート尺度で回答するよう依頼した（1:危険である～3:どちらでもない～5:安全である）。

その後、検索トピックに対する事前信念の回答に応じて、実験協力者を特定の実験条件に振り分けた。以下に詳細を記す。

- **biased(-)** 群: 「危険である」もしくは「やや危険である」と回答した実験協力者
- **critical** 群: 「どちらでもない」と回答した実験協力者
- **biased(+)** 群: 「安全である」もしくは「やや安全である」と回答した実験協力者

1 : <https://www.lancers.jp/>

次に、実験協力者には事前信念を強め、検索タスク時に確証バイアスを発生させるための事前情報を提示した。事前情報は検索タスクへの導入文、遺伝子組み換え食品に関する情報から構成される。導入文はすべての実験協力者群に共通である。一方、遺伝子組み換え食品に関する情報は、事前信念を強めるため、実験協力者群ごとに異なる。

検索タスクの導入文は以下のとおりである。

あなたは友人と一緒にスーパーに買い物に行きました。そこでセールになっていた「なたね油」を手にとると、「遺伝子組み換えなたねが含まれる可能性がある」と書かれた表示が目につきました。あなたは以前から「遺伝子組み換え食品」について少し気になっていたため、一緒にいた友人に遺伝子組み換え食品について意見を求めました。

実験協力者には、上記の導入文に加え、群に応じた事前信念を強めるための異なる情報を提示した。具体的には、**biased(-)**群に振り分けられた実験協力者には、「ヨーロッパでは遺伝子組み換え食品に対して厳しい規制を行っている」といったネガティブな情報を提示した。また、2分程度の遺伝子組み換え食品に反対する内容の動画²を閲覧するよう依頼した。これによって、**biased(-)**群には「遺伝子組み換え食品を摂取することは危険である」といった事前信念を強めることを期待した。

biased(+)群に振り分けられた実験協力者には、「厚生労働省は厳しい審査を行っており、多くの日本人は遺伝子組み換え食品を摂取している」といったポジティブな情報を提示した。また、2分程度の遺伝子組み換え食品に賛成する内容の動画³を閲覧するよう依頼した。これらの情報によって、**biased(+)**群には「遺伝子組み換え食品を摂取することは安全である」といった事前信念を強めることを期待した。

critical群には、**biased(-)(+)**群に提示したネガティブな情報とポジティブな情報の両方を提示した。これらの情報によって、**critical**群には「遺伝子組み換え食品を摂取することの安全性は賛否両論である」といった事前信念を強めることを期待した。なお、事前情報を確実に閲覧させるために、どの群に対しても動画閲覧時にその内容を100字程度でテキストでまとめさせた。

事前情報を閲覧した後、実験協力者はウェブ検索タスクに取り組んだ。実験協力者は、下記のような指示を読み、検索タスクを開始した。

下記の手順で遺伝子組み換え食品を摂取することの安全性について調べるタスクに取り組んでください。下記の「検索結果を表示」のボタンをクリックし、表示された検索結果リストおよびそのリンク先の情報を閲覧してください。「遺伝子組み換え食品を摂取することの安全性」について納得のいく結論が出せましたら、検索をやめて最終的なあなたの意見、その根拠を下記

フォームに入力してください。

実験協力者は検索を開始するボタンを押した後、表示された検索結果リスト（SERP）およびSERPからリンクされた文書を閲覧し、検索タスクを行った。本検索タスクでは、制限時間は設けなかった。実験協力者は自分の納得する情報が得られたら検索を終了し、検索タスクに対する回答（事後信念）を報告した。実験協力者は、遺伝子組み換え食品の摂取の安全性について、事前アンケートと同様、「1. 危険である」「2. やや危険である」「3. どちらでもない」「4. やや安全である」「5. 安全である」の選択肢から回答した。

検索タスク終了後、実験協力者はウェブ検索態度やヘルスリテラシー、属性情報に関する事後アンケートに回答した。検索態度に関するアンケートでは、ウェブ情報の品質判断に影響があるとされている下記項目[13]について、検索タスク中にどの程度重視したかを調査した。

- データや証拠の有無
- ウェブサイトに書かれた内容の新しさ
- ウェブサイトに書かれた内容のわかりやすさ
- ウェブサイトの見た目のきれいさ
- 情報提供者の専門知識の有無
- ウェブサイトに書かれた内容の詳細度

実験協力者は、ウェブ検索態度の質問に関して「1. まったく重視しなかった」「2. あまり重視しなかった」「3. そこそこ重視した」「4. かなり重視した」「5. 大いに重視した」の選択肢から回答した。

ウェブ上の健康情報を効果的に調査・利用するためのスキル「ヘルスリテラシー」の調査には、eHealth Literacy Scale (eHEALS) の日本語版[14]を利用した。実験協力者は8つの質問に関して、5段階のリッカート尺度（1:全くそう思わない～5:かなりそう思う）の選択肢から回答した。本研究では、eHEALSを実験協力者の批判的思考能力の尺度とした。属性情報に関するアンケートでは、性別、年齢、最終学歴を調査した。

3.4 検索結果リスト

本実験の検索タスクでは、Google 検索⁴やYahoo 検索⁵といった一般的なウェブ検索エンジンが返す検索結果リストを模したものを利用した。検索結果リストには、検索トピックに関してあらかじめ筆者が用意した検索結果を計30件表示した。図1は、実際に検索タスクで用いた検索結果画面である。

各検索結果は、実験開始前にGoogle検索にて、クエリ「遺伝子組み換え食品 “安全”」「遺伝子組み換え食品 “危険”」でウェブ検索を行うことで、それぞれ10件ずつ取得した。また、上記の両クエリに「site:ac.jp」もしくは「site:go.jp」を加えてウェブ検索を行うことで、大学または行政機関が発信する情報をそれぞれ5件ずつ取得した。この操作により、検索結果リストにおける情報の質を担保した。

その後、図2のように、取得した両クエリの検索結果を上位

2 : <https://www.youtube.com/watch?v=umXN64zIH-8>

3 : <https://www.youtube.com/watch?v=zMnX3qS6Dj4>

4 : <https://www.google.co.jp/>

5 : <https://www.yahoo.co.jp/>

SimpleSearch

遺伝子組み換え食品 安全性

遺伝子組み換え作物の日本における利用状況 | バイテク情報普及会

https://cbijapan.com/about_use/usage_situation_jp/

日本では、トウモロコシ、ダイズ、ナタネ、ワタなどの主要作物は海外からの輸入に大きく依存しており、その大部分が遺伝子組み換え品種と... 食品としてGM作物が多く使われているのは、コーン油、ダイズ油、ナタネ油、綿実油などの食用油、しょうゆ、コーンスターチ、コーンシロップなどです。... 厚生労働省「安全性審査の手続きを踏んだ旨の公表がなされた遺伝子組み換え食品及び添加物一覧」をもとにバイテク情報普及会取りまとめ...

遺伝子組み換え作物の現状と問題点 - 中野司朗レディスクリニック

https://www.clinica.jp/?post_type=keyword&p=2555

品種改良と違って、複数の生物の遺伝子を合体させて全く新しい生物を創造するのが遺伝子組換えです。例えば、ブタ... 現在日本で承認され、流通している遺伝子組換え作物は、ダイズ、トウモロコシ、ナタネ、ジャガイモ、綿、てん菜、アルファルファの7品目です。これは... また、ナタネ科の白菜などは自然交配の危険にさらされています。これらの...

「遺伝子組み換え」の安全性とは？なぜ賛否両論を巻き起こしているのか

<https://smartagi-jp.com/management/114>

大豆やトウモロコシは、遺伝子組み換え作物の事例としてよく知られている。また、そのほかの作物も、私たちがとって馴染みのあるものばかりだ。日本で遺伝子組み換え作物が使用されている食品...

「非遺伝子組換えナタネを守って!!」生協の食材宅配生活...

<https://seikatsuclub.coop/news/detail.html?NTC=00...>

生活クラブ生協連合会では、食用なたね油の原料として使用していたカナダ産ナタネにGMOが混入する危険性が高まったため、これを西オーストラリア州産に変更。これまで4回にも及ぶ「西蔵州NON-GM菜種ミッション」を実施し、組合員や職員、食用なたね油の生産者の(株)米澤製油... さらにオーストラリアの連邦政府は遺伝子組み換え(GM)ナタネの商業栽培を認めています。CBHと提携する農家がある西...

図 1 実験協力者に提示するウェブ検索結果画面

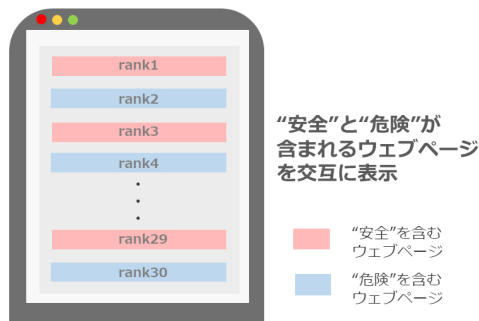


図 2 ウェブ検索結果一覧画面

から交互に配置することで、検索結果リストを作成した。このように配置することで、実験協力者の検索結果リスト閲覧時における、事前信念の影響を調べることが可能になる。なお、検索結果は一般的なウェブ検索の結果画面を模しているが、実験協力者がクエリの修正を行えないように設定した。

各検索結果をクリックすると、実験開始前にあらかじめアーカイブしたリンク先の文章が表示される。また、各文書の閲覧時間を計測できるよう、アーカイブしたウェブページに JavaScript のコードを埋め込んだ。なお、確実に文書閲覧時間を計測できるよう、文書中のハイパーリンクは無効化し、検索結果リストに表示された文書以外を閲覧できないように設定した。

3.5 計測データ

確証バイアスとウェブ検索行動の関係分析を行うために、検索タスク中の下記項目に関するデータを収集した。

- 検索結果一覧ページ閲覧時間（以下、SERP 閲覧時間）
- 最大ページ閲覧時間
- 検索セッション時間
- 検索結果のクリックスルー

表 1 各検索指標における condition 条件ごとの平均値と標準偏差を示す。

検索指標	condition		
	biased(-) 群	critical 群	biased(+) 群
検索セッション時間	446.6 (446.8)	437.0 (379.3)	269.7 (314.0)
SERP 閲覧時間	73.0 (86.6)	75.7 (63.4)	58.4 (82.9)
最大ページ閲覧時間	146.4 (118.5)	155.5 (144.9)	93.6 (58.5)
最大クリック深度	11.2 (9.2)	12.5 (9.9)	8.9 (9.4)
ページビュー数	5.0 (4.7)	5.3 (5.2)	5.1 (7.5)
信念の変化量	0.39 (1.15)	0.26 (0.99)	-0.35 (1.02)

SERP 閲覧時間は、実験協力者が SERP を閲覧した時間の合計である。最大ページ閲覧時間は、実験協力者が閲覧した SERP からリンクされたウェブページの中で、最大の閲覧時間である。検索セッション時間は、実験協力者のページ閲覧時間と SERP 閲覧時間の合計である。検索結果のクリックスルーは、実験協力者が SERP でクリックした検索結果の情報である。クリックスルー情報には、タイトル、概要文、URL、検索結果順位、信念の極性、公的機関が発信する情報かどうかが含まれる。これらの指標は、ウェブ検索行動のログを分析した White や Scott らの論文を参考に設定した [9] [15]。

4 結果

本章では、計 275 名の実験協力者に関するタスク中の行動データ、事前アンケート、事後アンケートの分析結果について述べる。本実験では、確証バイアスの有無（以下、condition 条件）と eヘルスリテラシースコア（以下、eHEALS）の 2 要因が検索閲覧行動・観点に及ぼす影響について分析を行った。condition 条件には、ネガティブな確証バイアスを有する群（biased(-) 群）、確証バイアスを有さない群（critical 群）、ポジティブな確証バイアスを有する群（biased(+) 群）の 3 水準を設定した。本研究では、critical 群の検索閲覧行動に対して、biased(-) 群と biased(+) 群にどのような違いがあるのかを分析した。

4.1 統計解析

本研究では、ベイズ統計モデリングの一つである一般化線形混合モデル (GLMM) を用いて、行動ログの分析を行った。GLMM は小規模なデータにおいて、頻度主義統計による方法よりも、正確に分析が行えるとされている [16]。また、検索行動のログは個人差が分散が大きいため、正確に分析することが難しい。GLMM はデータの分散を考慮したランダム効果を導入することで、過分散を抑えることができる。そのため、本研究では、分析手法としてベイズ統計モデリングを採用した。

ベイズ統計モデリングでは、事前分布と尤度関数を設定する。本研究ではすべての指標において、事前分布は弱情報事前分布に従うとした。Liu らによると、ウェブ検索時における時間はワイブル分布に従うとされている [17]。そのため、検索結果閲覧時間、最大ページ閲覧時間、検索セッション時間はワイブル分布に従うと仮定した。また、ページビュー数と最大クリック深度はポアソン分布に従うと仮定した。そして、信念の変化量

表 2 **critical** 群と比較したときの GLMM の結果. 数値は 90%HDI の中央値と区間を表している. 90%HDI 区間にゼロを含まない数値は太字で示す.

検索閲覧行動の指標	biased(-) 群			biased(+) 群		
	condition	eHEALS	交互作用	condition	eHEALS	交互作用
検索セッション時間	-1.01 [-1.96, -0.14]	-0.02 [-0.05, 0.01]	0.01 [-0.01, 0.04]	0.28 [-1.36, 1.91]	-0.02 [-0.05, 0.01]	-0.04 [-0.10, 0.03]
SERP 閲覧時間	-1.00 [-1.82, -0.11]	$4.58e^{-3}$ [-0.02, 0.03]	0.05 [0.01, 0.08]	0.89 [-0.57, 2.35]	$4.58e^{-3}$ [-0.02, 0.03]	-0.06 [-0.12, 0.02]
最大ページ閲覧時間	-0.54 [-1.37, 0.32]	-0.02 [-0.04, 0.01]	0.02 [-0.01, 0.05]	0.40 [-0.92, 1.78]	-0.02 [-0.04, 0.01]	-0.04 [-0.09, 0.02]
最大クリック深度	-1.02 [-1.97, -0.15]	$1.75e^{-3}$ [-0.03, 0.03]	0.04 [0.00, 0.07]	-0.20 [-1.76, 1.34]	$1.75e^{-3}$ [-0.03, 0.03]	-0.01 [-0.07, 0.05]
ページビュー数	-0.45 [-1.29, 0.37]	$9.92e^{-4}$ [-0.02, 0.03]	0.01 [-0.09, 0.04]	0.24 [-1.43, 1.79]	$9.92e^{-4}$ [-0.02, 0.03]	-0.02 [-0.09, 0.04]
ページビュー数 (-)	-0.14 [-1.15, 0.82]	$8.52e^{-3}$ [-0.02, 0.04]	$5.33e^{-3}$ [-0.03, 0.05]	-0.23 [-3.12, 2.94]	$8.52e^{-3}$ [-0.02, 0.04]	0.01 [-0.12, 0.14]
ページビュー数 (+)	-0.84 [-1.60, -0.09]	$-9.99e^{-3}$ [-0.03, 0.01]	0.03 [0.00, 0.06]	-0.12 [-1.55, 1.25]	$-9.99e^{-3}$ [-0.03, 0.01]	$4.58e^{-3}$ [-0.06, 0.05]
公的機関のページビュー数	-0.30 [-1.23, 0.64]	-0.01 [-0.04, 0.02]	$8.68e^{-3}$ [-0.03, 0.04]	0.39 [-2.12, 2.65]	-0.01 [-0.04, 0.02]	-0.03 [-0.12, 0.07]
信念の変化量	0.50 [-0.55, 1.44]	$9.03e^{-3}$ [-0.02, 0.04]	-0.34 [-1.80, 1.20]	0.39 [-2.12, 2.65]	$9.03e^{-3}$ [-0.02, 0.04]	-0.01 [-0.07, 0.05]

は正規分布に従うと仮定した.

GLMM において, condition 条件と eHEALS の 2 要因を主効果, 実験協力者をランダム効果とした. 具体的には, 下記の式に従い, R の brms パッケージを用いて分析を行った.

$Y \sim \text{condition} + \text{eHEALS} + \text{condition} : \text{eHEALS}$
 $+ (1 + \text{condition} + \text{eHEALS} + \text{condition} : \text{eHEALS} | \text{Participant})$

上記式において, Y は目的変数, condition は確証バイアスの有無を表し, eHEALS は e ヘルスリテラシスコアを意味する. また, (x|y) は y が x のランダム効果であることを意味する.

本研究では, condition 条件と eHEALS の 2 要因が及ぼす影響について調べるための指標として, 最高密度区間 (HDI) を用いた. HDI はパラメータの取りうる範囲を表したものである. この HDI の区間にゼロを含まない場合, パラメータが有効であることを意味する (頻度主義統計における帰無仮説を棄却することに相当). Kruschke の指摘にしたがい, 本研究では HDI の区間を 90% に設定した [18].

事後アンケートの分析は, ノンパラメトリックの分析手法を用いた.

4.2 検索セッション時間

実験協力者が検索・閲覧行動をどの程度慎重に行ったかを分析するために, 実験協力者間で検索セッションにかけた時間を比較した. 検索セッション時間は, SERP 閲覧時間および SERP からリンクされたウェブページの閲覧時間の合計値を用いた. 表 2 が示すように, **biased(-)** 群は, condition 条件の係数の 90%HDI にゼロが含まれていなかった (頻度主義統計における帰無仮説の棄却に相当). この結果より, **biased(-)** は **critical** 群よりも, 検索セッション時間が短い傾向にあるといえる. しかし, eHEALS, 交互作用の係数の 90%HDI にゼロが含まれていた (頻度主義統計における帰無仮説を棄却しないことに相

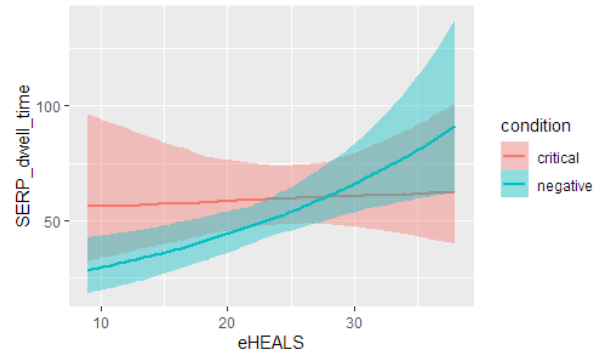


図 3 SERP 閲覧時間に対する condition 条件と eHEALS の効果の推定値. 赤色の線は **critical** 群, 青色の線は **biased(-)** 群を表す. 背景の色は信用区間を示す.

当). また, 表 2 が示すように, **biased(+)** 群は, condition 条件, eHEALS, 交互作用の係数の 90%HDI にゼロが含まれていた.

4.3 SERP 閲覧時間

実験協力者が情報を選ぶ際, どの程度慎重に走査したかを分析するため, SERP 閲覧時間を比較した. 表 2 が示すように, **biased(-)** 群は, condition 条件, 交互作用の係数の 90%HDI にゼロが含まれていなかった. しかし, eHEALS の係数の 90%HDI にゼロが含まれていた. 交互作用が確認されたため, 単純主効果分析を行った. 図 3 が示すように, **biased(-)** 群は **critical** 群よりも, eHEALS が低いと SERP 閲覧時間が短く, eHEALS が高いと SERP 閲覧時間が長い傾向にあった. また, 表 2 が示すように, **biased(+)** 群は, condition 条件, 交互作用の係数の 90%HDI にゼロが含まれていた.

4.4 最大ページ閲覧時間

実験協力者が SERP に掲載されたウェブページをどの程度慎重に閲覧したかを分析するために、実験協力者がタスク中に閲覧したウェブページの最大閲覧時間を比較した。表 2 が示すように、**biased(-)(+)** 群ともに、condition 条件、eHEALS、交互作用の係数の 90%HDI にゼロが含まれていた。この結果は、最大ページ閲覧時間に関して、確証バイアスの有無による影響がほとんどないことを示唆している。

4.5 ページビュー数

実験協力者が、検索トピックの真偽を客観的に確認する際、多数の証拠をどの程度収集しようとしたかを分析するために、タスク中に実験協力者が閲覧したウェブページ数を比較した。表 2 が示すように、**biased(-)(+)** 群ともに、condition 条件、eHEALS、交互作用の係数の 90%HDI にゼロが含まれていた。この結果は、ページビュー数に関して、確証バイアスの有無による影響がほとんどないことを示唆している。

詳細な分析を行うため、実験協力者が自身の信念と一致するページを、どの程度閲覧したかを分析した。なお、検索結果の概要文の中に、「危険」を含むものをネガティブ、「安全」を含むものをポジティブ、ネガティブなページのクリック数をページビュー数 (-)、ポジティブなページのクリック数をページビュー数 (+) とする。

表 2 が示すように、ページビュー数 (-) は **biased(-)(+)** 群ともに、condition 条件、eHEALS、交互作用の係数の 90%HDI にゼロが含まれていた。この結果は、ページビュー数 (-) に関して、確証バイアスの有無による影響がほとんどないことを示唆している。

表 2 に示すように、ページビュー数 (+) において、**biased(-)** 群は、condition 条件、交互作用の係数の 90%HDI にゼロが含まれていなかった。しかし、eHEALS の係数の 90%HDI にゼロが含まれていた。交互作用が確認されたため、単純主効果分析を行い、結果を図 5 に示す。図 5 が示すように、**biased(-)** 群は **critical** 群よりも、eHEALS が低いとページビュー数 (+) が少なく、eHEALS が高いとページビュー数 (+) が多い傾向にあった。また、表 2 に示すように、ページビュー数 (+) において、**biased(+)** 群は、condition 条件、eHEALS、交互作用の係数に 90%HDI にゼロが含まれていた。

実験協力者が公的機関の発信する信頼性の高い情報を、どの程度閲覧したかを分析した。表 2 に示すように、**biased(-)(+)** 群ともに、condition 条件、eHEALS、交互作用の係数の 90%HDI にゼロが含まれていた。この結果は、公的機関のページビュー数に関して、確証バイアスの有無による影響がほとんどないことを示唆している。

4.6 最大クリック深度

実験協力者が SERP をどの程度の深さまで走査したかを分析するために、実験協力者が SERP でクリックした検索結果の順位に着目し、最大の検索結果順位(最大クリック深度)を分析した。表 2 が示すように、**biased(-)** 群は、condition 条件、

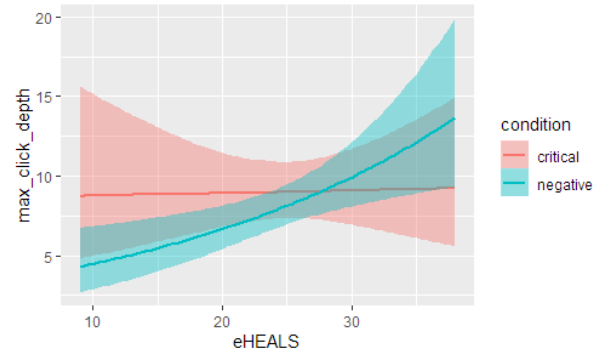


図 4 最大クリック深度に対する condition 条件と eHEALS の効果の推定値。赤色の線は **critical** 群、青色の線は **biased(-)** 群を表す。背景の色は信用区間を示す。

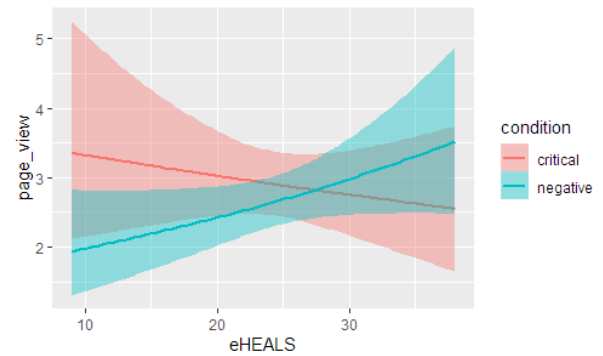


図 5 ページビュー数 (+) に対する condition 条件と eHEALS の効果の推定値。赤色の線は **critical** 群、青色の線は **biased(-)** 群を表す。背景の色は信用区間を示す。

交互作用の係数の 90%HDI にゼロが含まれていなかった。しかし、eHEALS の係数の 90%HDI にゼロが含まれていた。交互作用が確認されたため、単純主効果分析を行い、結果を図 4 に示す。図 4 が示すように、**biased(-)** 群は **critical** 群よりも、eHEALS が低いと最大クリック深度が浅く、eHEALS が高いと最大クリック深度が深くなる傾向にあった。また、表 2 が示すように、**biased(+)** 群は、condition 条件、eHEALS、交互作用の係数の 90%HDI にゼロが含まれていた。

4.7 信念の変化

検索タスクによって、事前信念がどの程度変化したかを分析するために、事後信念と事前信念の差を信念の変化量として分析を行った。表 2 が示すように、**biased(-)(+)** 群ともに、condition 条件、eHEALS、交互作用の係数の 90%HDI にゼロが含まれていた。この結果から、確証バイアスを有するユーザは、事前信念がほとんど変化しないといえる。

4.8 情報の精査観点

確証バイアスとウェブ検索・閲覧行動の関係を分析するために、行動データとは別に、実験協力者が検索タスク中に重視した情報精査観点の比較分析を行った。結果を表 3 に記す。表 3 が示すように、「データや証拠の有無」に関して、condition 条件による統計的有意差は確認されなかった (**biased(-)** 群: 3.01,

表 3 検索態度における、condition 条件ごとの平均値と標準偏差。統計的有意差が確認されたものは太字で示す。

情報精査観点	condition		
	biased(-) 群	critical 群	biased(+) 群
データや証拠の有無	3.01 (0.97)	3.04 (0.96)	2.87 (0.96)
内容の新しさ	2.47 (0.91)	2.58 (0.91)	2.48 (1.20)
内容のわかりやすさ	3.36 (0.87)	3.19 (0.80)	3.20 (0.75)
情報提供者の専門知識の有無	3.11 (1.10)	3.29 (0.92)	3.13 (1.09)
ウェブサイトの見た目	2.60 (1.02)	2.30 (0.87)	2.65 (0.91)
内容の詳細度	3.32 (0.82)	3.30 (0.78)	3.41 (0.72)

critical 群: 3.04, **biased(+)** 群: 2.87)。「内容の新しさ」に関して、condition 条件による統計的有意差は確認されなかった (**biased(-)** 群: 2.47, **critical** 群: 2.58, **biased(+)** 群: 2.48)。「内容のわかりやすさ」に関して、condition 条件による統計的有意差は確認されなかった (**biased(-)** 群: 3.36, **critical** 群: 3.19, **biased(+)** 群: 3.20)。「情報提供者の専門知識の有無」に関して、condition 条件による統計的有意差は確認されなかった (**biased(-)** 群: 3.11, **critical** 群: 3.29, **biased(+)** 群: 3.13)。一方、「ウェブサイトの見た目」に関して、condition 条件による統計的有意差が確認された (**biased(-)** 群: 2.60, **critical** 群: 2.30, **biased(+)** 群: 2.65, $p < .05$)。この結果から、確証バイアスを有するユーザは、そうではないユーザよりも、ウェブサイトの見た目を重要視しているといえる。「内容の詳細度」に関して、condition 条件による統計的有意差は確認されなかった (**biased(-)** 群: 3.32, **critical** 群: 3.30, **biased(+)** 群: 3.41)。

5 考 察

5.1 仮説検証

本研究では、仮説 **H1** を検証するため、SERP 閲覧時間と最大クリック深度を分析した。SERP 閲覧時間の結果から、**biased(-)** 群は **critical** 群よりも、eHEALS が低いほど時間が短く、eHEALS が高いほど時間が長い傾向にあった。また、最大クリック深度の結果から、**biased(-)** 群は **critical** 群よりも、eHEALS が低いほど上位の検索結果をクリックし、eHEALS が高いほど下位の検索結果をクリックする傾向にあった。一方、上記の検索指標において、**biased(+)** 群には違いが確認されなかった。

eHEALS はウェブ上の健康情報を批判的に精査するためのスキルである。そのため、eHEALS が低いとユーザは、どの情報を見るべきかに注意を払わない可能性が高い。上記の結果は、eHEALS の低い **biased(-)** 群が、**critical** 群よりも、どの情報を見るべきかに注意を払わない傾向を示している。このことから、eHEALS が低い **biased(-)** 群は、確証バイアスによる負の影響を受けやすいと考えられる。また、実験結果は、eHEALS の高い **biased(-)** 群が **critical** 群よりも、注意深く情報を閲覧していたことを示している。このことから、eHEALS の高い **biased(-)** 群は、検索トピックにネガティブな信念を持ちつつも、多くの意見を比較するため、情報を注意深く閲覧したと考えられる。そして、情報精査観点の結果から、確証バイアスを

有するユーザは、そうではないユーザよりもウェブサイトの見た目を重視していた。つまり、確証バイアスを有するユーザは、意思決定をする際に表層的な情報を重要視する傾向を示している。行動ログと情報精査観点の結果より、仮説 **H1** は支持されたといえる。

本研究では、仮説 **H2** を検証するため、最大ページ閲覧時間、検索セッション時間、ページビュー数を分析した。最大ページ閲覧時間において確証バイアスの有無による影響はみられなかった。一方、検索セッション時間において、**biased(-)** 群は、**critical** 群と比べて、時間が短い傾向が確認された。つまり、**biased(-)** 群は **critical** 群よりも、検索閲覧行動を注意深く行わない傾向にあると考えられる。また、ページビュー数の結果より、eHEALS の低い **biased(-)** 群は、**critical** 群よりも、自分の信念と一致しないウェブページの閲覧数が少ない傾向にあった。一方、eHEALS の高い **biased(-)** 群は、**critical** 群よりも、自分の信念と一致しないウェブページの閲覧数が多い傾向にあった。これらの結果は、SERP 閲覧時間・最大クリック深度と同様に、eHEALS の低い **biased(-)** 群は、確証バイアスの影響を受けやすく、異なる意見を閲覧しなかったことが原因であると考えられる。また、eHEALS の高い **biased(-)** 群は、情報を精査する能力があるため、自分の信念と異なる意見も閲覧したと考えられる。eHEALS の高い **biased(-)** 群は、異なる意見を多く閲覧したにも関わらず、自身の持つ信念はほとんど変化しなかった。このことから、検索トピックにネガティブな信念をもつユーザは、注意深く検索閲覧行動を行っても、確証バイアスの負の影響を取り除き、事前信念を変更することが難しいと考えられる。検索セッション時間とページビュー数の結果より、仮説 **H2** は部分的に支持されたといえる。

上記で述べたように、一部の検索指標において、**biased(-)** 群は、eHEALS の多寡に応じて結果が異なったことから、仮説 **H3** は部分的に支持されたといえる。

最後に、検索指標において、**biased(-)** 群にのみ違いが確認されたことについて述べる。これは、ネガティブな確証バイアスが、ポジティブな確証バイアスよりも、大きく影響したことが原因であると考えられる。Rozin らは、人間はポジティブな情報よりもネガティブな情報に影響を受けやすいことを報告している [19]。そのため、**biased(-)** 群は **biased(+)** 群よりも、事前信念による影響を受けやすく、結果的に **biased(-)** 群の行動ログに違いが生じたと考えられる。

5.2 改善点

本研究にて改善すべき点は 2 点ある。1 つ目は、検索トピックである。本研究では、健康分野から「遺伝子組み換え食品」を検索トピックとして扱った。しかし、健康分野以外で賛否両論の分かれる分野でユーザ実験を行い、分野ごとにおける確証バイアスの影響を調べる必要がある。

2 つ目は、SERP における情報の質である。今回、SERP には情報の質を担保するため、一定の割合で大学や公的機関の発信する情報を含めた。しかし、SERP の情報はポジティブな情報に信頼性の高いウェブページが偏っていた。今回、検索ト

ピックとして選んだ「遺伝子組み換え食品」は、日本では厚生労働省によって安全が確認されている食品である。そのため、ポジティブ側の検索結果に、厚生労働省といった国の公的機関が発信する情報が多く含まれていた。Liao らによると、自分の信念と一致しない情報であっても、情報に高い専門性であれば、クリックする傾向があるとされている [20]。つまり、ネガティブな確証バイアスを有するユーザは、ポジティブ側の情報に信頼性の高い情報が含まれていれば、信念の極性に関係なくクリックしてしまう可能性が高い。それゆえ、ユーザが検索結果を選んだ理由が、信念の影響か情報の信頼性の高さによるものかを正確に判断できたとはいえない。したがって、ネガティブ・ポジティブな情報の両方に、同程度の信頼性の高い情報を含めることが重要であるといえる。

6 おわりに

本研究では、確証バイアスを有するウェブ検索ユーザの検索行動を明らかにするため、クラウドソーシングを用いてオンライン実験を行った。確証バイアスを有する群とそうではない群に分けるため、検索トピックに関する印象を操作する事前情報を実験協力者に与え、その検索・閲覧行動のログを分析した。ユーザ実験の結果、検索トピックに対して否定的な信念を持つウェブ検索ユーザは、批判的に情報を精査するスキルが低い場合、検索結果一覧ページを閲覧する時間が短く、上位の検索結果をクリックし、肯定的な意見を閲覧しない傾向にあることが明らかになった。

謝 辞

本研究は JSPS 科研費 JP18KT0097, JP18H03243, JP18H03494, C18H032440 および課題設定による先導的な人文学・社会科学研究推進事業の助成を受けたものです。ここに記して謝意を表します。

文 献

- [1] Elizabeth Sillence, Pam Briggs, Lesley Fishwick, and Peter Harris. Trust and mistrust of online health sites. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '04, p. 663–670, New York, NY, USA, 2004. Association for Computing Machinery.
- [2] Huyen Le, Raven Maragh, Brian Ekdale, Andrew High, Timothy Havens, and Zubair Shafiq. Measuring political personalization of google news search. In *The World Wide Web Conference*, pp. 2957–2963. ACM, 2019.
- [3] Raymond S Nickerson. Confirmation bias: A ubiquitous phenomenon in many guises. *Review of general psychology*, Vol. 2, No. 2, pp. 175–220, 1998.
- [4] Ryen White. Beliefs and biases in web search. In *Proceedings of the 36th international ACM SIGIR conference on Research and development in information retrieval (SIGIR 2013)*, pp. 3–12. ACM, 2013.
- [5] Q Vera Liao and Wai-Tat Fu. Can you hear me now?: mitigating the echo chamber effect by source position indicators. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing (CSCW 2014)*, pp. 184–196. ACM, 2014.
- [6] Q. Vera Liao, Wai-Tat Fu, and Sri Shilpa Mamidi. It is all

- about perspective: An exploration of mitigating selective exposure with aspect indicators. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI 2015)*, p. 1439–1448, 2015.
- [7] Daniel Kahneman. *Thinking, fast and slow*. Macmillan, 2011.
- [8] Suppanut Pothirattanachai, Takehiro Yamamoto, Yusuke Yamamoto, and Masatoshi Yoshikawa. Analyzing the effects of document's opinion and credibility on search behaviors and belief dynamics. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management (CIKM 2019)*, p. 1653–1662, 2019.
- [9] Ryen W White and Dan Morris. Investigating the querying and browsing behavior of advanced search engine users. In *Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval (SIGIR 2007)*, pp. 255–262, 2007.
- [10] Carla Teixeira Lopes and Edgar Ramos. Studying how health literacy influences attention during online information seeking. In *Proceedings of the 2020 Conference on Human Information Interaction and Retrieval, CHIIR '20*, p. 283–291, New York, NY, USA, 2020. Association for Computing Machinery.
- [11] Takashi Kusumi, Rumi Hirayama, and Yoshihisa Kashima. Risk Perception and Risk Talk: The Case of the Fukushima Daiichi Nuclear Radiation Risk. *Risk Analysis*, Vol. 37, No. 12, pp. 2305–2320, 2017.
- [12] R.E. Petty and J.T. Cacioppo. The elaboration likelihood model of persuasion. *Advances in experimental social psychology*, Vol. 19, No. 1, pp. 123–205, 1986.
- [13] Yusuke Yamamoto, Takehiro Yamamoto, Hiroaki Ohshima, and Hiroshi Kawakami. Web access literacy scale to evaluate how critically users can browse and search for web information. In *Proceedings of the 10th ACM Conference on Web Science (WebSci 2018)*, p. 97–106, 2018.
- [14] 光武誠吾, 柴田愛, 石井香織, 岡崎勘造, 岡浩一朗. ehealth literacy scale (eheals) 日本語版の開発. 日本公衆衛生雑誌, Vol. 58, No. 5, pp. 361–371, 2011.
- [15] Scott Bateman, Jaime Teevan, and Ryen W. White. The search dashboard: How reflection and comparison impact search behavior. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2012)*, p. 1785–1794, 2012.
- [16] Matthew Kay, Gregory L. Nelson, and Eric B. Hekler. Researcher-centered design of statistics: Why bayesian statistics better fit the culture and incentives of hci. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, p. 4521–4532, New York, NY, USA, 2016. Association for Computing Machinery.
- [17] Chao Liu, Ryen W White, and Susan Dumais. Understanding web browsing behaviors through weibull analysis of dwell time. In *Proceedings of the 33rd international ACM SIGIR conference on Research and development in information retrieval*, pp. 379–386. ACM, 2010.
- [18] John Kruschke. *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan (English Edition)*. Academic Press, kindle edition edition, 11 2014.
- [19] Paul Rozin and Edward B. Royzman. Negativity bias, negativity dominance, and contagion. *Personality and Social Psychology Review*, Vol. 5, No. 4, pp. 296–320, 2001.
- [20] Q. Vera Liao and Wai-Tat Fu. Expert voices in echo chambers: Effects of source expertise indicators on exposure to diverse opinions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2014)*, p. 2745–2754, 2014.