

# 確証バイアスとウェブ検索行動の関係分析

鈴木 雅貴<sup>†</sup> 齊藤 史明<sup>†</sup> 山本 祐輔<sup>†</sup>

<sup>†</sup> 静岡大学情報学部 〒432-8011 静岡県浜松市中区城北3-5-1

E-mail: †{suzuki,saito,yamamoto}@design.inf.shizuola.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。クラウドソーシングで募集した89名のワーカーに対し、オンライン上で健康情報に関する検索タスクを依頼するオンラインユーザ実験を実施した。本実験では、検索タスク開始前に検索トピックに対する印象を操作する事前情報を提示することで、被験者の事前信念をコントロールし、検索タスク中の行動ログの分析を行った。実験の結果、確証バイアスを有するユーザは、ウェブ上の健康情報の批判的な情報閲覧能力である「ヘルスリテラシー」を有していてもそれをうまく活用することができず、検索結果の上位のみを閲覧し、手短かにウェブ検索を完了させる傾向にあることが明らかになった。

キーワード 情報信頼性, 情報検索, サーチエンジン, UI・UX

## 1 はじめに

近年、フェイクニュースをはじめとした誤情報がウェブ上に蔓延するようになってきている。ウェブ検索ユーザは、ウェブ情報が正しいか否かを見分けなければ、誤った情報を鵜呑みにしてしまう可能性がある。また、検索アルゴリズムによる最適化やレコメンドシステムによって、ウェブ検索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」が社会問題になりつつある [1]。フィルターバブルは、ウェブにおける誤情報の取得を助長する恐れがある。

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 [2]。例えば、ダイエット願望が強いユーザ X が、テレビ放送で商品 Y がダイエットに効果があるということを耳にし、商品 Y に期待を抱いたとする。このようなユーザ X が商品 Y の効果についてウェブ検索を行うと、ユーザ X は「商品 Y がダイエットに効果がある」と考えたいがために、たとえ商品 Y を批判する情報が提示されていたとしても、商品 Y を支持する情報を優先的に閲覧してしまう。結果的に、自分の信念に合致するという理由で誤まったウェブ情報を鵜呑みにしてしまう可能性がある。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 らによって明らかにされている [3]。

以上のように、確証バイアスは情報探索ユーザの情報精査行動を抑制しうる。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 [4]。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が活かす前提として、情報精査への動機付けが必要とされている [5]。確証バイアスは情報精査態度に影響を与えると考えられる。Kusumi らや Petty らの知見を踏まえると、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。

こうした中、情報検索やヒューマン・コンピュータ・インタラクション分野では、情報探索・閲覧時に生じる確証バイアスを軽減する方法について研究が行われつつある。例えば、Liao らは、ある観点で対立する情報を単純に両面呈示するのではなく、発信者の専門性と併せて提示することで、ユーザの確証バイアスに陥った状態を緩和できることを明らかにした [6] [7]。

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するか」といった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、健康トピックのウェブ検索に焦点を当て、確証バイアスを有するユーザが典型的に取り得るウェブ検索行動を明らかにする。

本研究では、被験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り、確証バイアスを有したユーザとそうではないユーザに分け、確証バイアスとウェブ検索行動の関係进行分析した。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：

確証バイアスを有するウェブ検索ユーザは、ものごとを多面的に判断せず、注意深い情報検索をすることが困難であると思われる。結果として、確証バイアスを有していないユーザよりも、検索結果の閲覧時間が短く、1 ページあたりのウェブページ閲覧時間が短く、検索タスクを始めてから終えるまでの時間が短く、総閲覧ウェブページ数が少なく、検索結果一覧の上位結

果を閲覧し手短かに検索行動を終える。このような行動は、特にウェブ情報を批判的に精査するスキルが元々高いユーザはほど顕著に表れる。

ユーザ実験の結果、ウェブ上の健康情報を批判的に精査するスキル「eヘルスリテラシー」スコアが高い被験者は、確証バイアスが発生していない状態では検索結果の下位にある情報を閲覧する一方で、確証バイアスが発生してしまうと、検索結果の上位を見て検索行動を終えてしまうことが明らかになった。

## 2 関連研究

### 2.1 ウェブ検索における認知バイアス

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 [8]。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 [3]。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い、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示した [9]。

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を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バイアスを軽減することを示した [7]。

### 2.2 ウェブ検索ユーザの検索・閲覧行動の分析

行動ログを用いたウェブ検索行動の分析に関する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 [10]。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。Yu らはユーザのウェブ検索・閲覧行動のログから知識量の状態、知識量の増加を予測するモデルの構築を行った [11]。

### 2.3 ウェブ検索ユーザの態度変容の促進

ウェブ検索ユーザの検索時の行動や態度を変容させるインタラクションに関する研究は数多く行われている。

Yamamoto らは、検索システムのクエリ補完/推薦時に批判的思考を想起させるクエリを挿入し、批判的情報検索を促進するシステム QUERY PRIMING を提案した [12]。このシステムを用いることによって、ウェブ検索ユーザはクエリ修正を頻繁に

行うようになり、データや証拠に基づいた情報を記したウェブサイトによく訪れることを明らかにした。

Scott らは、他のウェブ検索ユーザの検索・閲覧行動の履歴を可視化するシステム SEARCH DASHBOARD を提案した [13]。ユーザ実験の結果、当該システムを用いたウェブ検索ユーザは自身の検索行動を改善することが可能であることを明らかにした。

Ennals らは、賛否両論のある文章にハイライト表示をし、賛成意見と反対意見の両面を提示するシステム DISPUTE FINDER を提案した [14]。山本らは検索トピックに対する否定意見である「反証トピック」を提示するシステムを提案し、反証トピックの提示がウェブ検索ユーザの注意深い情報探索を促進することを明らかにした [15]。

## 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、検索トピックに対する印象を操作する事前情報を与えることで、被験者の事前信念を調整し、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、確証バイアスが発生するシナリオを割り当てられた群を biased 群、そうではない群を critical 群と呼称する。

### 3.1 被験者

クラウドソーシングサービス Lancers.jp<sup>1</sup>を用いて、100名の被験者を募った。タスクが完全に終了していない者、および何らかの理由で複数回タスクに取り組んでいる者のデータは除外した。本実験では、検索トピックに関する事前知識の有無が事前信念に影響する。そのため、タスク開始前に各被験者に検索トピックに関する知識量を確認し、検索トピックについて「まったく知識がない」もしくは「あまり知識がない」と答えた被験者のみを対象とした。最終的に、合計 89 名の被験者のデータを分析に用いた。

各被験者は確証バイアスの有無および検索トピックによって、特定の実験条件に割り当てられた。タスクの割り当て人数を表 1 に示す。後述するように、各被験者には実験条件に応じて異なる事前情報が与えられ、検索タスク開始前にトピックに対する事前信念の強さを 5 段階で回答させた。事前信念スコアの平均値は、biased 群が 2.56 (SD=0.67)、critical 群が 2.20 (SD=0.69) で、両群のスコアに統計的有意差が確認された ( $p < .05$ )。また、後述するヘルスリテラシースコアを biased 群と critical 群で比較したところ、両群間で統計的有意差は確認されなかった (biased=21.1; critical=21.7;  $p = 0.51$ )。タスクを終えた被験者に対しては、100 円の報酬を支払った。

### 3.2 検索タスク

検索トピックとして、健康食品として知られている「チア

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

表 1 被験者の割り当て

被験者群	検索トピック	
	チアシード	キヌア
biased	22	21
critical	22	24

シード」と「キヌア」の2つを設定した。これらトピックのいずれかに関して、被験者はそれがダイエットに有効であるかを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

本実験は以下の手順で実施した。

- (1) 被験者登録
- (2) 検索トピックに関する事前情報の提示
- (3) 事前アンケート
- (4) 検索タスク
- (5) 事後アンケート

まず、被験者は 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てもらった。その際、被験者に確認バイアスの有無条件、および検索トピックをランダムに割り当てた。

次に、被験者には検索タスクに取り組むための事前情報を提示した。以下は検索トピック「チアシード」の例である。

先日会社の健康診断を受けたあなたは、中性脂肪の数値が高く肥満傾向にあるという結果にショックを受けます。そこで 30 代にもなり、20 代の頃のような生活ではいけないと思ったあなたは、これをきっかけにダイエットすることを決意しました。そのため、まずは健康意識の高い同僚に体型を維持する秘訣を聞いてみると、「チアシードというスーパーフードを毎日食べることによって体型を維持している」ということを知ります。そして同僚からチアシードについて以下のような情報を教えてもらいました。

被験者には、上記事前情報に加えて、検索トピックに対する印象を操作するための事前情報を提示した。印象操作の方法は、[16] を参考にした。biased 群に対しては「チアシード（キヌア）がダイエットに効果がある」という信念（確認バイアス）を生じさせるポジティブな情報、critical 群に対しては「ネット上の健康情報の半分以上は信頼できない」といった対象トピックの真偽を慎重に精査するよう促す情報を与えた。

具体的には、biased 群に振り分けられた被験者には、「チアシード（キヌア）は脂肪を燃焼する成分のオメガ 3 脂肪酸を豊富に含む」といった情報を、図 1 のようにグラフとグラフの内容を要約した情報が添付された画像と共に紹介した。また、「海外の有名人が利用しており、流行になっている」といった情報を有名人の画像と共に紹介した。これらの情報によって、biased 群には「チアシード（キヌア）がダイエットに有効である」といった事前信念を持ってもらうことを期待した。

critical 群には「ウェブ上に存在する健康情報は半分以上が信頼できない情報である」といった情報を、図 2 のようにグラフとグラフの内容を要約した情報が添付された画像と共に与えた。また、「厚生労働省はウェブ上の健康情報に対して注意喚起を行っている」といった情報を、厚生労働省が発信している画像と共に紹介した。これらの情報によって、critical 群には「ウェブ上の健康情報は誤った情報が存在するため、精査する必要がある」といった事前信念を持ってもらうことを期待した。

事前情報を閲覧後、被験者は検索トピックに関する知識量、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、「チアシード（キヌア）について知っていますか?」という質問を課した。被験者は、「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。事前信念を問うアンケートでは、「チアシード（キヌア）はダイエットに有効であると思いますか?」という質問を課した。本質問に対しては、「1. まったく効果がない」「2. あまり効果がない」「3. そこそこ効果がある」「4. かなり効果がある」「5. 十分な効果がある」の選択肢から回答させた。

事前アンケート終了後、被験者はウェブ検索タスクに取り組んだ。被験者は、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「チアシード（キヌア）がダイエットに効果的か」を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。「チアシード（キヌア）がダイエットに効果的か」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。

被験者は検索を開始するボタンを押し、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、与えられたトピックがダイエットにどの程度有効かを調査した。本検索タスクでは、制限時間は設けなかった。被験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。検索タスクの回答は、事前アンケートと同様に、「1. まったく効果がない」「2. あまり効果がない」「3. そこそこ効果がある」「4. かなり効果がある」「5. 十分な効果がある」の選択肢から回答させた。

検索タスク終了後、被験者は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 [17] [18] 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

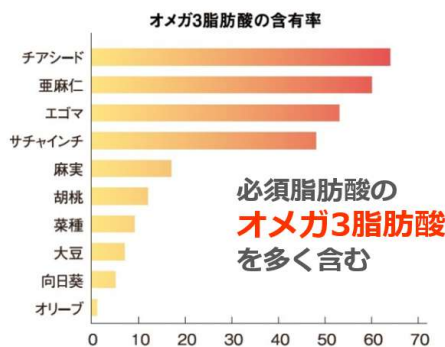


図1 biased群に与えた事前情報の例

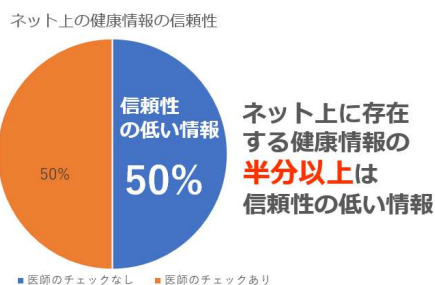


図2 critical群に与えた事前情報の例

● ウェブサイトに書かれた内容の詳細度

回答は5段階のリッカート尺度を用いた(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)。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 [19] を利用した。被験者は eHEALS に関する 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)にて回答した。具体的には「私は、インターネットでどのような健康情報サイトが利用できるかを知っている」、「私は、インターネット上のどこに役立つ健康情報サイトがあるか知っている」といった質問から構成されている。

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

本実験では、Google 検索<sup>2</sup>や Yahoo 検索<sup>3</sup>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を用いた。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 100 件表示した。具体的には、実験開始前に Google 検索にてクエリ「< 検索トピック > ダイエット “効果あり”」と「< 検索トピック > ダイエット “効果なし”」で検索を行い、図 3 のように得られた結果を上位から交互に 5 件ずつ配置することで、検索結果リストを作成した。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、被験者がクエリの修正を行えないように設定した。



図3 検索結果一覧画面

各検索結果をクリックすると、リンク先の文書が表示されるようにした。リンク先の文書は実験開始前にあらかじめアーカイブしたものを表示した。実験中の文書閲覧時間を計測できるように、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に、文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

検索行動の分析のため、検索タスク中の下記項目に関するデータを収集した。

- 検索結果閲覧時間
- 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

検索結果閲覧時間は、各検索トピックに関する検索結果リスト(以下、SERP)を、被験者が閲覧した時間を算出したものである。ページ閲覧時間は、SERP からリンクされたウェブページの閲覧時間である。検索セッション時間は、被験者のウェブページ閲覧時間と SERP 閲覧時間の合計である。検索結果のクリックスルーは、被験者が 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 らの論文を参考にした [10] [13]。

### 3.6 仮説

本研究では、確認バイアスを有するユーザは確認バイアスを有しないユーザと比較して、健康トピックに関するウェブ情報を注意深く検索しないと仮説を置く。具体的には、確認バイアスとウェブ検索行動の関係について、以下のような仮説を設定する：

**H1** 確認バイアスを有するユーザは、SERP 閲覧時間、平均ページ閲覧時間、検索セッション時間が短く、ページビュー数が少なく、最大クリック深度が浅くなる。

**H2** 確認バイアスを有するユーザは、ウェブ検索閲覧の観点が異なり、「データや証拠の有無」と「情報提供者の専門知識の有無」を重視しない。

また、第 1 章で述べたように、確認バイアスを有する場合、批判的思考態度が喚起されないため、たとえユーザが批判的に情報精査するスキルを有していても、それを有効活用すること

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

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

ができないと考えられる。そこで、本研究では、ウェブ上の健康トピックを精査するスキルであるヘルスリテラシーと確認バイアス、およびウェブ検索行動について、下記仮説を設定する。

**H3** 確認バイアスを有するユーザは、ヘルスリテラシーをうまく活用できず、SERP 閲覧時間、平均ページ閲覧時間、検索セッション時間が短く、ページビュー数が少なく、最大クリック深度が浅くなる。

## 4 結果

本章では計 89 名の被験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。本分析では、検索トピックである「チアシード」と「キヌア」はともに健康食品という括りであり、トピック間に差はないと仮定した。本実験ではヘルスリテラシースコアの程度（以下、eHEALS）と確認バイアスの有無（以下、condition）の 2 要因が検索閲覧行動・観点に及ぼす影響について分散分析を行った。eHEALS 要因には eHEALS の高い群（以下、high 群）と eHEALS の低い群（以下、low 群）の 2 水準を設定した。high 群は被験者全体の eHEALS の平均点（Mean=21.4）以上、low 群は平均点未満とした。condition 要因には確認バイアスを有する群（biased 群）と確認バイアスを有さない群（critical 群）の 2 水準を設定した。なお、実験で収集したデータには正規性が確認できなかったため、データに整列ランク変換 [20] を用いたノンパラメトリック分散分析を用いた。

### 4.1 検索結果閲覧時間

被験者が検索結果リスト（以下、SERP）をどの程度慎重に走査したかを分析するために、被験者間で SERP の閲覧時間を比較した。表 2 が示すように、eHEALS 要因、condition 要因、交互作用に統計的有意差は確認されなかった（eHEALS:  $p = .38$ ; condition:  $p = .89$ ; 交互作用:  $p = .25$ ）（ヘルスリテラシー high 群の平均: biased 群 72.7 vs. critical 群 63.0; ヘルスリテラシー low 群の平均: 69.0 vs. 62.1）。

### 4.2 検索セッション時間

被験者が検索・閲覧行動をどの程度慎重に行ったかを分析するために、被験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 の閲覧時間および 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、eHEALS 要因、condition 要因、交互作用に統計的有意差は確認されなかった（eHEALS:  $p = .31$ ; condition:  $p = .28$ ; 交互作用:  $p = .34$ ）（high 群の Mean: 293.7 vs. 248.1; low 群の Mean: 299.4 vs. 282.2）。

### 4.3 平均ページ閲覧時間

被験者が 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、被験者がタスク中に閲覧したウェブページの平均閲覧時間を算出し、被験者間で比較した。表 2 が示すように、eHEALS 要因、condition 要因、交互作用に統計的有意差は確認されなかった（eHEALS:  $p = .28$ ; condition:

$p = .13$ ; 交互作用:  $p = .52$ ）（high 群の Mean: 67.4 vs. 58.3; low 群の Mean: 51.9 vs. 54.5）。

### 4.4 ページビュー数

被験者が、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、タスク中に被験者が閲覧したウェブページ数を比較した。ページビュー数に関して、condition 要因と交互作用に統計的有意差は確認されなかった（condition:  $p = .99$ ; 交互作用:  $p = .28$ ）。一方、eHEALS 要因には統計的有意差が確認された（ $p < .05$ ）。表 2 が示すように、condition によらず low 群の被験者は high 群の被験者よりもページビュー数が多かった（high 群の Mean: 3.76 vs. 3.41; low 群の Mean: 4.89 vs. 5.24）。

### 4.5 最大クリック深度

被験者が SERP をどの程度の深さまで走査したかを分析するために、被験者が 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。最大クリック深度に関して、eHEALS 要因と condition 要因に統計的有意差は確認されなかった（eHEALS:  $p = .23$ ; condition:  $p = .97$ ）。一方、eHEALS 要因と condition 要因の交互作用には有意水準 10% で統計的有意差が確認された（ $p < .1$ ）。図 4、表 2 が示すように、high 群は biased 群のほうが critical 群に比べ、最大クリック深度の平均値が小さかった（15.6 vs. 19.6）。この結果は、ヘルスリテラシーを有する被験者に確認バイアスが発生すると、検索結果リストを下位の方まで見なくなる傾向にあることを示唆している。一方、low 群の場合は、biased 群の方が critical 群よりも最大クリック深度の平均値が大きかった（24.7 vs. 17.2）。これらの結果は、確認バイアスの有無がヘルスリテラシーの活用に一部影響を与えることを示唆している。

### 4.6 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、eHEALS 要因、condition 要因、交互作用に統計的有意差は確認されなかった（eHEALS:  $p = .49$ ; condition:  $p = .29$ ; 交互作用:  $p = .86$ ）（high 群の Mean: 0.57 vs. 0.71; low 群の Mean: 0.80 vs. 1.00）。

### 4.7 情報の精査観点

確認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、被験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。分析の結果、「データや証拠の有無」に関して、condition 要因、交互作用に統計的有意差は確認されなかった（condition:  $p = .93$ ; 交互作用:  $p = .65$ ）。一方、eHEALS 要因に統計的有意差が確認された（ $p < .001$ ）。表 3 が示すように、バイアスの有無によらず high 群の被験者は low 群の被験者よりも「データや証拠の有無」を重視していた（high 群の Mean: 2.90 vs. 3.04; low 群の Mean: 2.36 vs. 2.30）。

「内容の新しさ」に関して、condition 要因、交互作用に統

表 2 eHEALS と condition ごとにみる，検索閲覧行動の結果（括弧内は標準偏差を表す，\*\*\*: 有意水準 0.001, \*\*: 0.01, \*: 0.05, -: 0.1）.

検索指標	high		low		p-value		
	biased	critical	biased	critical	eHEALS	condition	交互作用
SERP 閲覧時間 (秒)	72.7 (92.4)	63.0 (68.9)	69.0 (66.3)	62.1 (65.6)	0.38	0.89	0.25
検索セッション時間 (秒)	293.7 (303.7)	248.1 (253.1)	299.4 (178.7)	282.2 (267.2)	0.31	0.28	0.34
平均ページ閲覧時間 (秒)	67.4 (51.9)	58.3 (63.3)	51.9 (35.6)	54.5 (63.5)	0.28	0.13	0.52
ページビュー数 (件)	3.76 (3.75)	3.41 (1.95)	4.89 (2.65)	5.24 (4.40)	*	0.99	0.28
最大クリック深度	15.6 (27.1)	19.6 (25.0)	24.7 (29.1)	17.2 (21.3)	0.23	0.97	.
信念の変化量	0.57 (0.79)	0.71 (0.81)	0.80 (0.77)	1.00 (0.91)	0.49	0.29	0.86

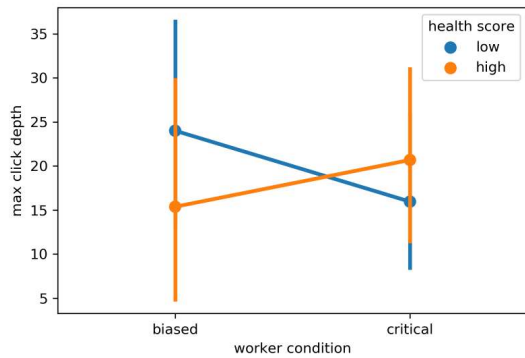


図 4 最大クリック深度（プロットは平均値，エラーバーは信頼区間を示す）

計的有意差は確認されなかった（condition:  $p = .53$ ; 交互作用:  $p = .34$ ）。一方，eHEALS 要因に統計的有意差が確認された ( $p < .001$ )。表 3 が示すように，バイアスの有無によらず high 群の被験者は low 群の被験者よりも「内容の新しさ」を重視していた（high 群の Mean: 2.52 vs. 2.42; low 群の Mean: 1.68 vs. 1.80）。

「内容のわかりやすさ」に関して，eHEALS 要因，condition 要因，交互作用に統計的有意差は確認されなかった（eHEALS:  $p = .43$ ; condition:  $p = .73$ ; 交互作用:  $p = .19$ ）（high 群の Mean: 3.33 vs. 3.23; low 群の Mean: 3.00 vs. 3.30）。

「情報提供者の専門知識の有無」に関して，condition 要因，交互作用に統計的有意差は確認されなかった（condition:  $p = .50$ ; 交互作用:  $p = .84$ ）。一方，eHEALS 要因に統計的有意差が確認された ( $p < .001$ )。表 3 が示すように，バイアスの有無によらず high 群の被験者は low 群の被験者よりも「情報提供者の専門知識の有無」を重視していた（high 群の Mean: 2.86 vs. 2.73; low 群の Mean: 2.09 vs. 1.90）。

「ウェブサイトの見た目」に関して，eHEALS 要因，condition 要因，交互作用に統計的有意差は確認されなかった（condition:  $p = .43$ ; 交互作用:  $p = .44$ ）（high 群の Mean: 2.67 vs. 2.31; low 群の Mean: 2.41 vs. 2.85）。

「内容の詳細度」に関して，eHEALS 要因，condition 要因，交互作用に統計的有意差は確認されなかった（eHEALS:  $p = .85$ ; condition:  $p = .35$ ; 交互作用:  $p = .53$ ）（high 群の Mean: 3.33 vs. 3.19; low 群の Mean: 3.00 vs. 2.85）。

## 5 考 察

### 5.1 事前情報による影響

biased 群と critical 群の事前信念スコアが示しているように，本実験で提示した事前情報によって，被験者の事前信念は調整することにある程度成功したと考えられる。すなわち，biased 群は「対象トピックがダイエットに有効である」という確証バイアスを持ち，critical 群は「対象トピックがダイエットに有効であるかは慎重に見定める必要がある」という事前信念を持たせることにある程度成功したと考えられる。

事前に立てた仮説では，biased 群は確証バイアスを有するため，対象トピックがダイエットに有効であることを支持する情報を適当に集め，比較的手短かにウェブ検索タスクを終えると予想した。一方で，critical 群は対象トピックのダイエットに対する有効性を慎重に調査するために，比較的最長い時間，かつ多くの情報を求めてウェブ検索タスクに取り組むことを予想した。しかし，予想に反して critical 群と biased 群の検索行動に差はみられなかった。これは critical 群に与えた「ウェブ上の健康情報の半分以上は信頼できない情報である」といった情報精査を促す事前情報によって，検索トピックに対して懐疑的になり，真剣に検索タスクに取り組まなかった可能性が考えられる。

### 5.2 確証バイアスとヘルスリテラシーの関係

#### 5.2.1 検索閲覧行動

検索閲覧行動について，eHEALS 要因と condition 要因によって分析を行ったところ，SERP 閲覧時間，検索セッション時間および平均ページ閲覧時間に差はみられなかった。我々は，eHEALS の高い被験者は，本来有する健康情報に対する批判的な情報閲覧能力により情報閲覧を注意深く行うため，SERP 閲覧時間および検索セッション時間が長くなると予想した。一方，biased 群は注意深い情報探索能力をすることが困難であるため，これらの時間が短くなると予想した。しかし，予想を支持する結果を得られなかった。

ページビュー数に関しては，eHEALS の低い群ほど多くのページを閲覧した。我々は，eHEALS の高い群ほど本来有する批判的な情報閲覧能力により，客観的な証拠を得ようとするため，多くのページを閲覧すると予想した。しかし，予想を支持する結果は得られなかった。これは，eHEALS の高い被験者ほど怪しげな健康食品といったものに対して懐疑的であり，多くのウェブページを見る必要がないと判断したためと考えられる。



表3 eHEALS と condition ごとにみる、情報精査観点の結果（括弧内は標準偏差を表す、\*\*\*: 有意水準 0.001, \*\*: 0.01, \*: 0.05, -: 0.1）。

観点	high		low		p-value		
	biased	critical	biased	critical	eHEALS	condition	交互作用
データや証拠の有無	2.90 (0.83)	3.04 (0.77)	2.36 (0.90)	2.30 (0.80)	**	0.93	0.65
内容の新しさ	2.52 (0.98)	2.42 (0.86)	1.68 (1.09)	1.80 (0.83)	***	0.53	0.34
内容のわかりやすさ	3.33 (0.80)	3.23 (0.65)	3.00 (0.69)	3.30 (0.73)	0.27	0.73	0.19
情報提供者の専門知識の有無	2.86 (1.01)	2.73 (0.87)	2.09 (1.06)	1.90 (0.79)	***	0.50	0.84
ウェブサイトの見た目	2.67 (1.06)	2.31 (1.01)	2.41 (1.01)	2.45 (1.01)	0.75	0.43	0.44
内容の詳細度	3.33 (0.66)	3.19 (0.69)	3.00 (0.62)	2.85 (0.75)	0.85	0.35	0.53

最大クリック深度の分析結果は、eHEALS が高く確認バイアスを有する被験者ほど、上位の検索結果を閲覧することを示している。一般的にウェブ検索ユーザは、検索結果順位の上位をクリックしやすいということが知られている [21]。そして、健康情報に関する批判的な情報閲覧能力を有する eHEALS の高い群は、より下位の検索結果を閲覧すると予想した。しかし、確認バイアスを有する被験者は eHEALS が高いほど最大クリック深度が浅くなる傾向にあった。この結果から、確認バイアスはヘルスリテラシーの活用にも負の影響を与える傾向にあると明らかになった。すなわち、部分的に仮説 H3 を支持したといえる。また、eHEALS が低く確認バイアスを有していない被験者ほど上位の検索結果を閲覧した。これは eHEALS が低い被験者ほど、健康情報に関する批判的な情報閲覧能力が低いため、「ウェブ上に存在する健康情報は信頼できない情報が多い」という事前情報を信じやすい傾向にあったと思われる。結果として、検索トピックに対して懐疑的になり、下位のほうまで見る必要がないと判断したと考える。

以上の結果から、**H1** と **H3** は部分的に支持されたと考える。

### 5.2.2 情報精査観点

情報精査観点について、eHEALS 要因と condition 要因によって分析を行ったところ、「データや証拠の有無」、「内容の新しさ」および「情報提供者の専門知識の有無」は eHEALS の高い群ほど重視していた。これらは批判的に情報を閲覧する際に必要となる項目である。そのため、eHEALS の高い被験者ほどこれらの観点を重視することは予想された。しかし、確認バイアスの有無と交互作用は確認されなかった。これは情報精査観点に関して、確認バイアスは負の影響を与えないことを示唆している。「内容のわかりやすさ」は biased 群、critical 群ともに他の観点到比べて重視する傾向にあった。一般的にウェブサイトに書かれた内容がわかりやすいほどクリックされやすいことが知られている [22] が、critical 群にはわかりやすさを重視しないことを予想した。しかし、予想を支持する結果は得られなかった。「ウェブサイトの見た目」については biased 群、critical 群ともに差がなかった。一般的にデザインが洗練されていると信じやすいという傾向にあると知られている [23] が、critical 群にはウェブサイトの見た目を重視しないと予想した。しかし、予想を支持する結果は得られなかった。「内容の詳細度」は biased 群、critical 群ともに他の観点到比べて重視する傾向にあった。biased 群は注意深い閲覧行動をすることが難しく、詳細度を重視しないと予想した。しかし、予想を支持する結果は得られなかった。以上の結果から、被験者が情報探索をする際

の観点到確認バイアスによる影響が及ばないと明らかになった。すなわち、仮説 H2 は支持されなかった。

### 5.3 今後の課題

今回行った実験では、あまり知られていないが一般に興味関心が高いと思われる健康食品の「チアシード」と「キヌア」のダイエットの効果を検証するウェブ検索タスクを実施した。しかし、巷にあふれる健康食品といった情報は怪しげなものも多い。それゆえ、特に検索トピックに懐疑的な事前情報を与えられた critical 群の被験者は、検索をする前からトピックに対して疑念を持ってしまい、検索タスクを真剣に行わなかった可能性がある。結果として、興味のある検索トピックについて、確認バイアスを有するウェブ検索ユーザとそうでないユーザの検索行動を比較するという、当初の目的が達成できなかった可能性がある。確認バイアスがウェブ検索行動に与える影響を厳密に分析するためには、少なくとも被験者間の検索へのモチベーションを均質化することが必要である。そのためには、被験者のトピックに対する関心を惹きつけるシナリオを作る、または検索トピックに対する事前の関心度を測る必要がある。

## 6 おわりに

本研究では、確認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。確認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被験者に与え、その検索・閲覧行動のログを分析した。ユーザ実験の結果、仮説に反して、確認バイアスの有無によってウェブ検索行動の傾向およびウェブ検索時に重視する観点到差はみられなかった。一方、確認バイアスの有無と健康情報に対する批判的な情報閲覧能力であるヘルスリテラシースコアを要因とする分析を行った結果、確認バイアスを有するユーザはヘルスリテラシースコアが高いほど、上位の検索結果を見て検索行動を終える傾向にあることが明らかになった。一般的に検索結果の上位結果を優先的に閲覧する傾向にあることが知られている [24] が、ヘルスリテラシースコアの高い被験者は、健康情報に対する批判的な情報閲覧能力が高いため、より下位の検索結果を閲覧すると考えた。しかし、ヘルスリテラシースコアの高い被験者は、確認バイアスを有すると、上位の検索結果を閲覧する傾向にあった。このことから、確認バイアスを有するユーザは、健康情報の批判的な情報閲覧能力「ヘルスリテラシー」

をうまく活用することができないということが明らかになった。

今後は、検索トピックや事前情報の提示の仕方を変えて実験を再度行い、確認バイアスを有するユーザのウェブ検索行動の理解を深める必要がある。

## 謝 辞

本研究は JSPS 科研費 JP16H02906, JP18KT0097, JP18H03243, JP18H03244, JP18H03494, および課題設定による先導的人文学・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] 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.
- [2] Raymond S Nickerson. Confirmation bias: A ubiquitous phenomenon in many guises. *Review of general psychology*, Vol. 2, No. 2, pp. 175–220, 1998.
- [3] Ryen White. Beliefs and biases in web search. In *Proceedings of the 36th international ACM SIGIR conference on Research and development in information retrieval*, pp. 3–12. ACM, 2013.
- [4] 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.
- [5] 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.
- [6] 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*, pp. 184–196. ACM, 2014.
- [7] 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 '15, p. 1439–1448, New York, NY, USA, 2015. Association for Computing Machinery.
- [8] ダニエル・カーネマン. 村井章子 (訳): ファスト & スロー (上). 早川書房, 2012.
- [9] Suppanut Pothirattanachaikul, 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 '19, p. 1653–1662, New York, NY, USA, 2019. Association for Computing Machinery.
- [10] 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*, pp. 255–262. ACM, 2007.
- [11] Ran Yu, Ujwal Gadiraaju, Peter Holtz, Markus Rokicki, Philipp Kemkes, and Stefan Dietze. Predicting user knowledge gain in informational search sessions. In *The 41st International ACM SIGIR Conference on Research Development in Information Retrieval*, SIGIR '18, p. 75–84, New York, NY, USA, 2018. Association for Computing Machinery.
- [12] Yusuke Yamamoto and Takehiro Yamamoto. Query priming for promoting critical thinking in web search. In *Proceedings of the 2018 Conference on Human Information Interaction Retrieval*, CHIIR '18, p. 12–21, New York, NY, USA, 2018. Association for Computing Machinery.
- [13] 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 '12, p. 1785–1794, New York, NY, USA, 2012. Association for Computing Machinery.
- [14] Rob Ennals, Beth Trushkowsky, and John Mark Agosta. Highlighting disputed claims on the web. In *Proceedings of the 19th International Conference on World Wide Web*, WWW '10, p. 341–350, New York, NY, USA, 2010. Association for Computing Machinery.
- [15] 山本祐輔, 嶋田敏. 検索トピックに対する反証示唆がウェブ検索ユーザの情報精査態度に与える影響. 人工知能学会論文誌, Vol. 32, No. 1, pp. WII-L.1–12, 2017.
- [16] Ha-Kyung Kong, Zhicheng Liu, and Karrie Karahalios. Trust and recall of information across varying degrees of title-visualization misalignment. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, New York, NY, USA, 2019. Association for Computing Machinery.
- [17] 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.
- [18] 山本祐輔, 山本岳洋, 大島裕明, 川上浩司. ウェブアクセスリテラシー尺度の開発. 情報処理学会論文誌データベース (TOD), Vol. 12, No. 1, pp. 24–37, 2019.
- [19] 光武誠吾, 柴田愛, 石井香織, 岡崎勘造, 岡浩一郎. ehealth literacy scale (ehealth) 日本語版の開発. 日本公衆衛生雑誌, Vol. 58, No. 5, pp. 361–371, 2011.
- [20] Jacob O. Wobbrock, Leah Findlater, Darren Gergle, and James J. Higgins. The aligned rank transform for nonparametric factorial analyses using only anova procedures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, p. 143–146, New York, NY, USA, 2011. Association for Computing Machinery.
- [21] Thorsten Joachims, Laura Granka, Bing Pan, Helene Hembrooke, Filip Radlinski, and Geri Gay. Evaluating the accuracy of implicit feedback from clicks and query reformulations in web search. *ACM Trans. Inf. Syst.*, Vol. 25, No. 2, p. 7–es, April 2007.
- [22] Charles L. A. Clarke, Eugene Agichtein, Susan Dumais, and Ryen W. White. The influence of caption features on click-through patterns in web search. In *Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '07, p. 135–142, New York, NY, USA, 2007. Association for Computing Machinery.
- [23] Gitte Lindgaard, Cathy Dudek, Devjani Sen, Livia Sumegi, and Patrick Noonan. An exploration of relations between visual appeal, trustworthiness and perceived usability of homepages. *ACM Trans. Comput.-Hum. Interact.*, Vol. 18, No. 1, May 2011.
- [24] Yisong Yue, Rajan Patel, and Hein Roehrig. Beyond position bias: Examining result attractiveness as a source of presentation bias in clickthrough data. In *Proceedings of the 19th International Conference on World Wide Web*, WWW '10, p. 1011–1018, New York, NY, USA, 2010. Association for Computing Machinery.