

賛否両論ある検索トピックにおけるウェブページ評価と意見形成の調査

萩原 諒† 山本 岳洋††

† 兵庫県立大学 社会情報科学部 〒651-2197 兵庫県神戸市西区学園西町 8-2-1

†† 兵庫県立大学 大学院情報科学研究科 〒651-2197 兵庫県神戸市西区学園西町 8-2-1

E-mail: †fa19e063@u-hyogo.ac.jp, ††t.yamamoto@sis.u-hyogo.ac.jp

あらまし 本研究では、賛否両論あるトピックにおいてどのようなウェブページを参考しているのか、そして最終的にどのように自身の意見を形成しているのかの2つを調査した。具体的には、「小学生に宿題は必要か」と「子供に新型コロナウイルスのワクチンを打つべきか」の2つのトピックを用いてユーザ実験を行い、12名の実験参加者から得られた意見形成のデータ、ウェブページ評価のデータを分析した。ウェブページ評価の分析では、高評価のウェブページ評価データに先行研究をもとに作成したラベルを付与した。ラベル付けの結果、宿題タスクの場合、多角的な視点からの情報や意見が記載されているウェブページや、実験結果のような根拠に基づいた情報が記載されているウェブページを参考に行っていることがわかった。ワクチンタスクの場合、多くの実験参加者が実験結果のような根拠に基づいた情報が記載されているウェブページを参考に行っていることがわかった。また、意見形成の分析では、意見形成理由の自由記述を手で分類した。その結果、損害最小型、メリット重視型、デメリット重視型、信用性重視型、多数決型、単一ページ確信型、事前信念先行型の7つの型があることがわかった。

キーワード 情報検索、賛否両論あるトピック、検索行動分析

1 はじめに

今日、ウェブ検索を用いて様々な重要な意見形成が行われている。たとえば、重要な意見形成には、「緑茶ががん予防に効果があるのか」などの医療情報に関するものや、「ベジタリアンになるべきか」などの賛否両論あるトピックに関するものなどが挙げられる。2020年のMozによる報告によると、回答者の46%がGoogleを使用して、しばしばまたはとても頻繁に医療の重要な決定を行っていることが明らかになっている[11]。

このような重要な意見形成を目的とした検索においては、検索者は種々の危険性や問題を意識しながら情報を収集し自身の意見を形成する必要がある、これまでに検索結果の偏りや誤りが指摘されている[8][17]。また、自身の考えから生じる無意識のバイアスは意見形成や検索行動に影響を与えるとされている[12][17]。

そのため、重要な意見形成において、どのようなウェブページを参考に行っているのか、そして最終的にどのように自身の意見を形成しているのかを明らかにすることは、ウェブ検索が検索者に与える影響やより良い意見形成を可能とするシステム実現のために重要である。健康情報については、どのような情報が信頼できるのか[16]、また、複数の情報源から情報が得られた際、どのように得られた情報を集約して意見形成をしているのか[20]に関する研究がある。

健康情報に関する検索においては、前述したように信憑性の分析や意見形成の分析など様々な研究がある一方で、賛否両論あるトピックに関する検索についてはそのような研究が十分になされていない。ここで本研究では、賛否両論あるトピックを真偽が明らかになっていなく賛成意見と反対意見が存在し議論

になっているトピックと定義する。たとえば、「小学生に宿題は必要か」や「子供に新型コロナウイルスのワクチンを打つべきか」などが挙げられる。このような賛否両論あるトピックは、答えが決まっていなく様々な情報から慎重な意見形成が必要だと考えられる。そのため、賛否両論あるトピックにおける検索結果や検索行動に関する研究[1][8]、クエリからの検出に関する研究[3][9]などがなされている。これらの研究は、賛否両論あるトピックを扱った検索システムの分析を行っている。しかし、賛否両論あるトピックにおける検索者とのインタラクションを考慮した分析を行っている研究は少ない。

そこで本研究では、賛否両論ある検索トピックにおいて、どのようなウェブページを参考に行っているのか、最終的にどのように自身の意見を形成しているのかを分析する。本研究でのリサーチクエストは以下の通りである。

RQ1: 賛否両論あるトピックにおいて、どのようなウェブページを参考に行っているのか。

RQ2: 賛否両論あるトピックにおいて、最終的にどのように意見形成を行っているのか。

これらのリサーチクエストを明らかにするためにユーザ実験を行った。ユーザ実験では、事前に用意した「小学生に宿題は必要か」、「子供に新型コロナウイルスのワクチンを打つべきか」の2つの賛否両論あるトピックのどちらかを自動で実験参加者に割り当て、著者らが用意した検索システムを用いて検索タスクを実施した。検索タスク終了後、検索したトピックに対する意見とその意見を形成した理由、検索タスク中に閲覧した全てのウェブページに対して、参考になったかどうかを尋ねるタスク後アンケートとインタビューを行った。アンケートとインタビューの結果を分析することで、賛否両論あるトピックにおいて、どのようなウェブページを参考に行っているのか、そ

して最終的にどのように自身の意見を形成しているのかを明らかにできると考えられる。

RQ1 を明らかにするために、タスク後アンケートで得られたウェブページ評価のデータとインタビュー内容を用いて分析した。具体的には、閲覧したウェブページの中で、高評価のウェブページに対し先行研究 [16] [20] をもとに作成したラベルを付与した。ラベル付けの結果、宿題タスクの場合、実験結果のような根拠に基づいた情報が記載されているウェブページや、多角的な視点からの情報や意見が記載されているウェブページが参考になることがわかった。また、ワクチンタスクの場合、実験結果のような根拠に基づいた情報が記載されているウェブページを多くの人が参考にしていることがわかった。

RQ2 を明らかにするために、タスク後アンケートで得られた意見形成のデータとインタビュー内容を用いて分析した。具体的には、実験参加者が回答した意見形成理由の自由記述から様々な意見形成の型に分類した。その結果、損害最小型、メリット重視型、デメリット重視型、信用性重視型、多数決型、単一ページ確信型、事前信念先行型の 7 つの型があることがわかった。

2 関連研究

本節では関連研究について述べる。まず賛否両論あるトピックに関する関連研究について述べ、次に検索行動の分析に関する関連研究、最後に意見形成の分析に関する関連研究を述べる。

2.1 賛否両論あるトピック

賛否両論あるトピックにおける研究は様々なされている。具体的には、投入されたクエリから賛否両論あるトピックかどうかを特定する研究 [3] [9] や、検索結果やユーザの検索行動に関する研究 [1] [6] [8] などがある。たとえば、Karl らは Google のクエリ補完機能を確認することで検索クエリが賛否両論あるトピックかどうかを特定することが可能であることを明らかにしている [9]。Gizem らは賛否両論あるトピックに関連するクエリの検索結果はバイアスがかかる傾向があることを明らかにしている [8]。

また、賛否両論あるトピックにおいて賛成意見と反対意見の両方を入手できる検索システムの研究がなされている [4] [5] [15]。Christian らは賛否両論あるトピックを検索した際に、そのトピックに対する賛成意見と反対意見を可視化する ArgumenText という検索システムを開発し、そのシステムの有用性を明らかにした [15]。

一方で、どのようなウェブページを参考にしているのかやどのように意見形成を行っているのかを明らかにしている研究は少ない。本研究では、この 2 つに着目して分析を行う。

2.2 検索行動の分析

検索者のウェブ検索行動を明らかにする研究がなされている [14] [18] [19]。Yamamoto らはウェブ検索を通して信頼できる情報を獲得するために必要な行動に対する心がけを検証態度と定義し、検証態度に影響を与える要因と、検証態度と実

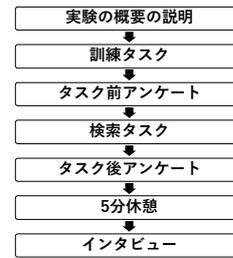


図 1 ユーザ実験の手順。

際のウェブ検索行動との関係について明らかにしている [19]。Suppanut らは検索前に持っていた意見と相反する意見を含んだ文書を開覧すると、検索プロセスにより労力をかけ、検索後に検索前の考えへと変える可能性が高くなることを明らかにした [14]。

2.3 意見形成の分析

検索者が健康情報を検索した際、どのように意見を形成しているのかに関する研究がなされている。浜島らは健康情報について複数の情報源から情報が得られたとき、それらをどのように集約して自身の意見を形成しているのか、そしてウェブページに対する信憑性判断が意見形成にどのような影響を与えているのかを明らかにしている [20]。White らは、健康情報における yes-no 型のクエリに対し、検索者は SERP に否定的な結果を記載しているウェブページはあまり閲覧せず、肯定的な結果を記載しているウェブページを閲覧する傾向にあることを明らかにしている [17]。

また、ウェブ検索を行い意見を形成する際検索エンジン操作効果 (SEME) があることがわかっている [2] [13]。SEME とは、検索結果を意図的にどちらかの意見に偏らせることで検索者の意見形成に影響を与える効果のことである。Pogacar らは検索結果を正しい情報で偏らせた場合は正しい意見形成になりやすくなり、誤った情報で偏らせた場合誤った意見形成になりやすいことを明らかにした [13]。

3 実験計画

本節ではまず、ユーザ実験の手順について説明し、その後実験参加者と検索タスクにおけるトピックの選定について詳細に述べる。最後に、ユーザ実験で行うタスク前アンケート、検索タスク、タスク後アンケート、インタビューについて詳細に述べる。

3.1 実験手順

ユーザ実験は図 1 に示した流れで行った。具体的にはまず、実験参加者に対して検索タスクの概要や収集するデータなどについての説明を行った。その際、実験中に収集するデータを学術目的に利用することを説明し、実験に同意した実験参加者には同意書に署名をしてもらった。次に、検索タスクを行う前に、著者らが用意した検索システムに慣れるための訓練タスク

を行った。訓練タスクでは、「ベジタリアンになるべきか」というトピックについて検索してもらった。訓練タスク終了後、実験参加者にタスク前アンケートを回答してもらった。タスク前アンケートでは、検索トピックに対する興味関心の有無、事前知識の有無、事前の考えの3項目を調査した。そして、タスク前アンケート回答後に検索タスクを行ってもらい、検索タスク終了後、タスク後アンケートを回答してもらった。タスク後アンケートでは、検索後の検索トピックに対する実験参加者の意見とその理由について回答してもらい、その後閲覧した全てのウェブページに対し、参考になったかどうかの評価とその理由を回答してもらった。最後に、タスク後アンケートの内容をより詳細に知るために、口頭でのインタビューを行いユーザ実験を終了した。

3.2 実験参加者

2022年11月15日から2022年12月5日にかけて、著者が所属する大学の学部生を対象に実験参加者を募集し、12名(男性6名、女性6名)の実験参加者が集まった。実験の開始から終了までかかった時間は、一番短い人で約45分、長い人で1時間30分であった。実験を終了した12名の実験参加者には1455円を支払った。

3.3 検索タスクにおけるトピックの選定

本実験で用いたトピックはProCon.org¹に掲載されているものの中から選んだ。ProCon.orgとは賛否両論あるトピックをまとめたウェブサイトである。実験に一般性を持たせるため2つのトピックを用いて実験を行った。具体的には、「小学生に宿題は必要か」と「子供に新型コロナウイルスのワクチンを打つべきか」を選んだ。以後、「小学生に宿題は必要か」を検索するタスクを宿題タスク、「子供に新型コロナウイルスのワクチンを打つべきか」を検索するタスクをワクチンタスクと呼ぶ。

3.4 タスク前アンケート

検索タスクを行う前に、検索トピックについての興味関心と事前知識の有無、そして検索タスクの事前の意見の3つの項目を調査するタスク前アンケートを行った。

宿題タスクの場合、興味関心を問う質問では、「あなたは小学生の宿題は必要かどうかについてどの程度興味関心がありますか。」と質問した。質問の回答には、5段階のリッカート尺度を用いた(1:まったく関心がない, 2:あまり関心がない, 3:どちらともいえない, 4:やや関心がある, 5:かなり関心がある)。事前知識を問う質問では、「あなたは小学生の宿題は必要かどうかについてどの程度知識があると思いますか。」と質問し、質問の回答には5段階のリッカート尺度を用いた(1:まったく知識がない, 2:あまり知識がない, 3:どちらともいえない, 4:やや知識がある, 5:かなり知識がある)。最後に、実験参加者の事前の意見を問う質問では、「小学生の宿題は必要だと思いますか。」と質問した。質問の回答には、5段階のリッカート尺度を用いた(1:まったくそう思わない, 2:あまりそう思

1: <https://www.procon.org/>



図2 ユーザ実験で用いた検索システムのインタフェースのスクリーンショット。(a) 検索タスク、(b) 検索結果のインタフェース

わない, 3:どちらともいえない, 4:ややそう思う, 5:かなりそう思う)。ワクチンタスクでも同様の質問を行い、検索タスク前の実験参加者の意見を調査した。

3.5 検索タスク

検索タスクを開始するにあたり、実験参加者には検索タスクの状況設定や内容の説明などが書かれたタスク文を読んでもらい、その後検索タスクに取り組んでもらった。たとえば宿題タスクの場合は、以下のタスク文を読んでもらった。

この検索タスクでは、対象を小学生に限定します。また、宿題とは教師から義務的に課された課題と定義します。

あなたは最近テレビで、宿題について議論している番組を見ました。今日、AIドリルという新しい宿題方法が広がる一方、宿題自体を廃止する学校もあります。

そこで、このタスクではウェブ検索を用いて、小学生に宿題は必要かどうかを自由に調べてください。

また、検索タスクに取り組む時間は最大30分とし、実験参加者が満足した検索ができたと感じた場合には30分より前に検索を止められるようにした。もし、28分経っても検索を終えない場合は「十分時間が経ったので後2分ほどで検索を終えてください」と伝えた。30分という時間を設定した理由は、十分に検索してほしいからである。検索タスクにおいては、著者が用意した検索システムのみを使用してもらった。検索システムではMicrosoft Bing Web Search API²を用いて、入力した

2: <https://www.microsoft.com/en-us/bing/apis/bing-web-search-api>

表 1 タスク後アンケートの内容。

ID	質問内容
Q1	閲覧したウェブページをどの程度見ていましたか。
Q2	閲覧したウェブページでは宿題は必要かどうかと言及していましたか。
Q3	閲覧したウェブページの内容は以下のどれにあてはまっていますか。
Q4	閲覧したウェブページは、自分の意見を形成するうえでどの程度参考になりましたか。
Q5	Q4 の項目でなぜそのような選択をしましたか、理由を教えてください。

検索クエリに対応する検索結果を表示した。実際にユーザ実験で用いた検索システムのインタフェースを図 2 に示す。

3.6 タスク後アンケート

検索タスクを行った後に、検索後の検索トピックに対する実験参加者の意見と検索タスク中に閲覧したウェブページの評価を調査するタスク後アンケートを行った。

3.6.1 検索トピックに対する実験参加者の意見

実験参加者には、検索したトピックについてどのような意見を持ったかを選択肢の中から選び、その意見になった理由を自由記述欄に記述してもらった。

具体的には宿題タスクの場合まず、検索トピックについてどのような意見を持ったかを問う質問では、「小学生の宿題は必要だと思いませんか。」と質問する。回答は「そう思う」、「そう思わない」、「今回の検索ではわからなかった」の 3 つの選択肢の中から選んでもらい、その選択肢を選んだ理由を自由記述欄に記述してもらった。検索トピックについての意見を調査した後、検索トピックの意見が反対にならなかった理由についても質問した。たとえば、小学生の宿題は必要であると回答した実験参加者には「なぜ、小学生の宿題は必要でないと回答しなかったのか思いつく限りの理由を記入してください。」と質問した。ワクチンタスクの場合も同様に質問した。これらのアンケートで実験参加者の検索後の意見を調査した。

3.6.2 ウェブページの評価

検索者がどのようなウェブページを参考にしているのかを調査するために、実験参加者が検索タスク中に閲覧した全てのウェブページに対して評価をしてもらった。実験参加者に質問した項目は表 1 に示した Q1 から Q5 の 5 つである。Q1 は「閲覧したウェブページをどの程度見ていましたか。」と尋ね、回答には 5 段階リッカート尺度を用いた (1: まったく見ていない, 2: あまり見ていない, 3: どちらともいえない, 4: やや見ていた, 5: かなり見ていた)。「まったく見ていない」と回答した実験参加者には、Q2, Q3, Q4 の質問で「Q1 でまったく見ていないを選んだ方はこちらを選択してください。」を回答してもらい、Q5 の質問には「Q1 でまったく見ていないを選んだ方は「特になし」と記入してください。」を回答してもらった。Q2 は「閲覧したウェブページでは小学生の宿題は必要かどうかと言及していましたか。」と尋ね、回答には「言及していた」と「言及していなかった」の 2 つの選択肢から実験参加者に選んでもらった。Q3 は「閲覧したウェブページの内容は以下のどれにあてはまっていますか。」と尋ねた。回答には、「検索タスクに対する賛成意見を記述していた」「検索タスクに対する

反対意見を記述していた」「検索タスクに対する賛成意見・反対意見を記述していた」「関係のない内容だった」の 4 つの選択肢から実験参加者に選んでもらった。Q4 は「閲覧したウェブページは、自分の意見を形成するうえでどの程度参考になりましたか。」と尋ね、回答には 7 段階リッカート尺度を用いた (1: 非常に参考にならなかった, 2: かなり参考にならなかった, 3: やや参考にならなかった, 4: どちらともいえない, 5: やや参考になった, 6: かなり参考になった, 7: 非常に参考になった)。最後に Q5 は「Q4 の項目でなぜそのような選択をしましたか、理由を教えてください。」と尋ね自由記述で回答してもらった。これらのアンケートでどのようなウェブページが参考になるのかを調査した。

3.7 インタビュー

実験参加者がタスク後アンケートを終え 5 分間の休憩をとった後に、タスク後アンケートで回答した内容についてより詳細に調査するために半構造化インタビューを行った。インタビュー時は、実験参加者が回答したタスク後アンケートと閲覧したウェブページの評価を見ながら質問し、インタビューの内容は全て録音した。あらかじめ 2 つの質問内容を用意し、その質問の回答に応じて質問内容を変えた。1 つ目は意見形成に理由について質問した。具体的には、「様々なウェブページを見たと思うのですが、なぜその意見に決めたのかを教えてください。」と質問した。2 つ目は検索を終えた理由について質問した。具体的には、「検索を終えたと思うのですが、どのような情報を得て検索を終えたのですか。」と質問した。30 分間検索をした実験参加者に対しては、「長い間検索をしてもらったと思うのですが、満足のいく検索はできましたか。」と質問した。

また、タスク後アンケートで回答したウェブページの評価とその理由について質問した。たとえば、閲覧したページの 1 つにかなり参考になったと回答し、その理由の自由記述欄に「メリットとデメリットが書かれたページだったから」と書かれていた場合、「なぜメリットデメリットの両面の内容が書かれていれば参考になるのですか。」と質問した。

それぞれの質問に対し、実験参加者の回答に応じて追加の質問をした。

4 分 析

本節では、1 節で述べた 2 つのリサーチクエスチョンを明らかにするためにウェブページ評価の分析、意見形成の分析の 2 つに分けて分析した結果について詳細に述べる。

表2 ウェブページ評価の自由記述に対するラベル。

ラベル名	定義
専門性	著者は情報を適切に扱うことができる十分な知識を持っているかどうか
権威性	情報を提供している組織が社会的に信用されているかどうか
客観性	情報が個人の考えや商業的な意図がない事実に基づいているかどうか
人気	情報が人気の情報(様々なサイトに書かれている情報)であるかどうか
詳細性	情報が詳細な情報であるかどうか
両面提示	情報がメリットとデメリット両方について述べているかどうか
了解性	内容がわかりやすいかどうか
意見の有無	メリットやデメリットが明確に述べられているかどうか
実例	情報が実体験に基づいているかどうか
視点の有無	様々な当事者視点の情報であるかどうか
適合性	求めている情報と合っているかどうか
その他	上記のラベル以外の基準によって評価されている

表3 ウェブページ評価の結果。

	宿題タスク	ワクチンタスク	両タスクの合計
専門性	0	1	1
権威性	0	2	2
客観性	4	12	16
人気	1	1	2
詳細性	3	3	6
両面提示	3	1	4
了解性	0	1	1
意見の有無	1	1	2
実例	2	2	4
視点の有無	5	2	7
適合性	26	20	46
その他	9	0	9
計	54	46	100

4.1 ウェブページ評価の分析

本節では、検索者が参考にするウェブページの要因を分析した結果について述べる。まず、どのようなデータを用いたのかと分析方法について述べ、その後、分析した結果について詳細に述べる。

4.1.1 分析方法

ウェブページの評価に対する分析を行うため、タスク後アンケートで実験参加者が回答した、閲覧したウェブページ評価の理由に対する自由記述とインタビュー内容を対象として分析を行った。具体的には、ウェブページの評価において実験参加者が「やや信頼できる」、「かなり信頼できる」、「非常に信頼できる」の3つのどれかと評価したウェブページに対してラベル付けを行った。ウェブページの評価に対するラベル付けではウェブページごとに集計した。

今回用いるラベルは、2つの先行研究をもとに著者らが作成した[16][20]。作成したラベルとその定義を表2に示す。ラベル付けは著者ら2名で議論を行いながら付与し、ラベルの定義を満たす内容の記述があるかどうかで2値分類した。1つの文に対して複数のラベルが付与できる場合には複数のラベルを付与した。いずれのラベルの定義にも満たさない文に対してはど

のラベルも付与していない。また、分析の際「適合性」のラベルについては分析対象から除いた。その理由は、検索時自分が求めている情報が手に入るウェブページが参考になるのは当然だと考えたからである。

4.1.2 分析結果

表3にウェブページ評価におけるラベル付けの結果を示す。表3より、宿題タスクの場合、視点の有無が一番評価され、続いて客観性が評価された。また、詳細性や両面提示なども評価された。閲覧したウェブページが両面提示されていて参考になると回答した実験参加者にインタビューの際理由を尋ねると、以下のような発話がみられた。

“一人の人がメリットもデメリットも書いている方がいろんなことを考えている気がするから。”(ID5)

また、閲覧したウェブページに客観性があり参考になると回答した実験参加者にインタビューの際理由を尋ねると、以下のような発話がみられた。

“研究されている方が信憑性がある感じがするから。”(ID3)

これらの結果から、宿題タスクの場合、多角的な視点からの情報や意見が記載されているウェブページや、実験結果のような根拠に基づいた情報が記載されているウェブページが参考になることがわかった。

ワクチンタスクの場合、客観性が実験参加者の多くに評価され、続いて詳細性が評価された。ワクチンタスクのトピックは健康情報の性質も含んでいるため、実験結果などの根拠が記載されているウェブページを参考にする結果になったと考えられる。健康情報における意見形成において客観性のある情報が参考になることは、浜島らが明らかにしており同様の傾向が見られた[20]。閲覧したウェブページに客観性があり参考になると回答した実験参加者にインタビューの際理由を尋ねると、以下のような発話がみられた。

表 4 意見形成の分類結果.

	宿題タスク	ワクチンタスク
損害最小型	2	0
メリット重視型	0	3
デメリット重視型	0	2
信用性重視型	2	0
多数決型	1	0
単一ページ確信型	1	0
事前信念先行型	1	0

“メディアや記事は主観が入ってくるので、信用できるのは研究された方のデータだと思っています。” (ID7)

これらの結果から、ワクチンタスクの場合、実験結果のような根拠に基づいた情報を記載しているウェブページを多くの方が参考にしていることがわかった。

4.2 意見形成の分析

本節では、検索者が得られた情報からどのように意見を形成しているのかに関する分析結果について述べる。まず、どのようなデータを用いたのかと分析方法について述べ、その後、分析結果について詳細に述べる。

4.2.1 分析方法

意見形成に対する分析を行うため、タスク後アンケートで実験参加者が回答した意見形成の理由に対する自由記述とインタビュー内容を対象にして分析を行った。具体的には、実験参加者が検索後に決定した意見形成の理由における自由記述を、著者らが議論しながら人手で様々な型に分類した。また、同じ実験参加者が複数の意見形成の型を持つこともある。たとえば、ID1の実験参加者の意見形成において、「損害最小型」と「事前信念先行型」の二つの型を持つこともあり得る。

4.2.2 分析結果

分析の結果、意見形成には少なくとも以下の7つの型がみられた。「多数決型」、「単一ページ確信型」、「事前信念先行型」は浜島らの研究で明らかになった意見形成の型であり、本実験でも同じ結果が得られた [20]。

- 損害最小型
- メリット重視型
- デメリット重視型
- 信用性重視型
- 多数決型
- 単一ページ確信型
- 事前信念先行型

これらの7つの型について以下で詳細に述べる。

a) 損害最小型

損害最小型とは、マイナスな結果になり過ぎないように結論付ける意見形成である。損害最小型に分類された実験参加者が意見形成の理由で回答した自由記述を以下に示す。

“多すぎる宿題を出すのはよくないが、全く出さないと学業

に支障が出てきそうだった。” (ID3)

このように、利益を重視するのではなく損害をできるだけ小さくするような意見形成を行う実験参加者がみられた。

b) メリット重視型

メリット重視型とは、与えられたトピックに対するメリットとデメリットを比較した際、メリットの方がデメリットよりも大きいと考える意見形成である。メリット重視型に分類された実験参加者が意見形成の理由で回答した自由記述を以下に示す。

“新型コロナウイルスのワクチンでの副反応よりコロナウイルスにかかったり重症化した時のほうが症状が重かった。” (ID9)

このように、デメリットは感じつつもメリットがはるかに大きいと判断し意見形成を行う実験参加者がみられた。

c) デメリット重視型

デメリット重視型とは、与えられたトピックに対するメリットとデメリットを比較した際、デメリットの方がメリットの方が大きいと考える意見形成である。デメリット重視型に分類された実験参加者が意見形成の理由で回答した自由記述を以下に示す。

“2200人を対象としたファイザーの実験で、コロナワクチンを打つことで、90%の感染リスクを防ぐことができると明らかになったと「ワクチンを打つべきか」と検索すると出てきました。この結果からだと、打ったほうが良いのかなとも思いましたが、日本の子ども(10代以下)のコロナ重症率が0%台ということを踏まえると、子どもにワクチンを打つメリットがあまり感じられませんでした。” (ID7)

このように、デメリットと比べてメリットが小さいと判断し意見形成を行う実験参加者がみられた。

d) 信用性重視型

信用性重視型とは、信用性の高いウェブページを参考にする意見形成である。信用性とは、「発信者が最も妥当と考える主張を伝えようとする意図に対する確信度」と定義されている [7]。信用性重視型に分類された実験参加者が意見形成の理由で回答した自由記述を以下に示す。

“反対的な意見を書いているサイトも閲覧したが、多くの記事がクーパー教授の「多すぎる宿題はデメリットが多い」という研究結果を歪曲し、宿題を課すべきではないと書かれているだけだった。” (ID11)

このように、公正なウェブページの情報を重視して意見形成を行う実験参加者がみられた。

e) 多数決型

多数決型とは、閲覧した多くのウェブページで主張されている意見を重視する意見形成である。多数決型に分類された実験参加者が意見形成の理由で回答した自由記述を以下に示す。

“宿題には勉強の習慣化という意図が含まれているという記事が多くあり、その意見に自分も賛成しているから。” (ID10)

このように、閲覧したウェブページの情報の中でより多くのウェブページに記述されている情報を重視して意見形成を行う実験参加者がみられた。

f) 単一ページ確信型

単一ページ確信型とは、最も参考のできるウェブページの情報を重視する意見形成である。単一ページ確信型に分類された実験参加者が意見形成の理由で回答した自由記述を以下に示す。

“宿題についての研究を行っているハリス・クーパー教授の研究で、宿題を行うことによる多くのメリットが述べられているから。” (ID11)

このように、閲覧したウェブページの中で最も参考のできるウェブページを重視して意見形成を行う実験参加者がみられた。

g) 事前信念先行型

事前信念先行型とは、検索タスクに対する事前の考えを重視する意見形成である。事前信念先行型に分類された実験参加者が意見形成の理由でインタビュー時に発話した内容を以下に示す。

“塾講師をしているので、経験則的なもので宿題は必要だと感じていた。” (ID5)

このように、検索前の意見と同じ意見が書かれているウェブページの情報を重視して意見形成を行う実験参加者がみられた。

5 議 論

本節では、前節の分析で得られた知見に対する考察と示唆について述べ、その後、本研究で行った分析の限界点について述べる。

5.1 考察と示唆

4.1 節のウェブページ評価の分析結果から、宿題タスクの場合、様々な立場を含んでいるため多角的な視点からの情報や意見が評価されたと考えられる。「小学生に宿題は必要か」のトピックには、小学生の立場、親の立場、学校の先生の立場、塾の立場、社会的立場など、様々な立場における利害が存在する。そのため、少数の立場からの情報や意見ではなく様々な観点からの情報が評価されたと考えられる。また、ワクチンタスクの場合、「子供に新型コロナウイルスのワクチンを打つべきか」というトピックは健康情報に近い性質を持っている。健康情報は人命に直結するため、先行研究にあるように、個人の意見ではなく確固たる根拠に基づいた情報が参考になったと考えられる [20]。

4.2 節の意見形成の分析の結果から、多数決型、単一ページ確信型、事前信念先行型は健康情報における意見形成と同様の

結果が得られた [20]。また、今回扱った賛否両論あるトピックは賛成意見と反対意見が存在するため、両意見を比較するような意見形成の型（損害最小型、メリット重視型、デメリット重視型）が見られたと考えられる。

4.1 節のウェブページ評価の分析と 4.2 節の意見形成の分析の結果から、意見形成の型は様々であり、それぞれの型の検索者に対し参考になるウェブページを提示すれば慎重な意見形成が促せると考えられる。たとえば、信用性重視型の検索者は信用性が低いウェブページをあまり参考にせず意見形成を行う傾向がある。これは、自分の意見と反対の意見が記載されているウェブページの信用性が高い場合、その意見に決定する可能性が高いということ意味する。つまり、信用性を重視する検索者は、閲覧するウェブページがどの程度信用できるかによって意見が大きく変わる可能性があることを示している。そのため、信用性重視型の検索者にはより多くの信用性の高いウェブページを提示することで、賛成意見と反対意見をそれぞれ吟味しながら自身の意見を形成することを促せると考えられる。以上のように、それぞれの型の検索者に対しその検索者が参考にできると考えるウェブページを多く見せることで慎重な意見形成が促せると考えられる。

5.2 限界点

本研究での 3 つの限界点について述べる。まず本研究では、実験参加者を筆者と同じ情報学部の実生から募集したため実験参加者に偏りがあると考えられる。Hölscher らの研究で、検索に熟練した検索者は一般の検索者と比べて、複雑なクエリを作成し自分の求めている情報を得やすいということがわかっている [10]。この研究から、情報学部の生徒は検索に熟練した検索者である可能性があり、一般的な検索者よりも適合性の高い検索結果を見つけられたと考えられる。そのため、検索者の年代や検索の専門性を考慮したユーザ実験を行えばより詳細なデータが得られると考えられる。

次に、今回のユーザ実験では、「子供に新型コロナウイルスのワクチンを打つべきか」と「小学生に宿題は必要か」の 2 つのトピックを扱った。しかし、トピックには様々な性質を含んでいる。たとえば、今回扱った「子供に新型コロナウイルスのワクチンを打つべきか」のトピックは、賛否両論の性質と健康情報の性質を含んでいる。このように、トピックには様々な性質を含んでいるため、本実験では 2 つのトピックを選んだが、「学生服は必要か」や「安楽死は合法か」などの賛否両論あるトピックの数を増やして実験を行うことで、本実験では得られなかった参考になるウェブページや意見形成の型を見つけることができると考えられる。

最後に、本研究ではユーザ実験で実験参加者が回答したアンケートとインタビュー内容を用いて分析を行った。しかし、実験参加者の検索行動のログデータを用いて分析を行うことで、検索行動の観点からの支援が可能であると考えられる。たとえば、多様なクエリを実験参加者が投入できていなかった場合、検索クエリの推薦によって多様な情報を得ることができ、意見形成を支援することができると考えられる。また、片方の意見

が記載されたウェブページを多数閲覧していた場合、もう一方の意見が記載されたウェブページを提示することで多様な情報に触れるための支援をすることができると考えられる。今後は、実験参加者の検索行動に着目した分析を行うことを考えている。

6 ま と め

本研究では、賛否両論あるトピックにおいてどのようなウェブページを参考にしているのか、そして最終的にどのように自身の意見を形成しているのかをユーザ実験を行い調査した。まず、ウェブページ評価の分析では、実験参加者が回答したウェブページ評価のデータに、先行研究をもとに作成したラベルを著者らで議論しながら付与した。ラベル付けの結果、宿題タスクの場合、多角的な視点からの情報や意見が記載されているウェブページや、実験結果のような根拠に基づいた情報が記載されているウェブページを参考にしていることがわかった。また、ワクチンタスクの場合、多くの実験参加者が実験結果のような根拠に基づいた情報が記載されているウェブページを参考にしていることがわかった。次に、意見形成の分析では、タスク後アンケートで実験参加者が回答した意見形成理由の自由記述を用いて人手で様々な型に分類した。分類した結果、損害最小型、メリット重視型、デメリット重視型、信用性重視型、多数決型、単一ページ確信型、事前信念先行型の7つの型があることがわかった。

本研究の知見を活かすことにより、賛否両論あるトピックを検索した際、検索者の意見形成の型を考慮し、参考にされやすいウェブページを提示する検索システムの有用性が示唆されたと考えられる。

謝辞 本研究は JSPS 科学研究費助成事業 JP21H03774, JP21H03775, による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] Yamen Ajjour, Pavel Braslavski, Alexander Bondarenko, and Benno Stein. Identifying argumentative questions in web search logs. *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 2393–2399, 2022.
- [2] Ahmed Allam, Peter Schulz, and Kent Nakamoto. The impact of search engine selection and sorting criteria on vaccination beliefs and attitudes: Two experiments manipulating google output. *Journal of medical Internet research*, Vol. 16, No. e100, 2014.
- [3] Sergiu Chelaru, Ismail Sengor Altinogvde, Stefan Siersdorfer, and Wolfgang Nejdl. Analyzing, detecting, and exploiting sentiment in web queries. *ACM Transactions on the Web*, Vol. 8, No. 1, pp. 1–28, 2013.
- [4] Sihao Chen, Daniel Khashabi, Chris Callison-Burch, and Dan Roth. Perspectroscope: A window to the world of diverse perspectives. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pp. 129–134, 2019.
- [5] Tim Draws, Oana Inel, Nava Tintarev, Christian Baden, and Benjamin Timmermans. Comprehensive viewpoint representations for a deeper understanding of user interactions with debated topics. In *Proceedings of the ACM SIGIR Conference on Human Information Interaction and Retrieval*, pp. 135–145, 2022.

- [6] Tim Draws, Nava Tintarev, Ujwal Gadiraju, Alessandro Bozzon, and Benjamin Timmermans. This is not what we ordered: Exploring why biased search result rankings affect user attitudes on debated topics. In *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 295–305, 2021.
- [7] Brian. J. Fogg. *Persuasive technology: Using computers to change what we think and do*. Morgan Kaufmann, 2002.
- [8] Gizem Gezici, Aldo Lipani, Yücel Saygin, and Emine Yilmaz. Evaluation metrics for measus results. *Information Retrieval Journal*, Vol. 24, pp. 85–113, 2021.
- [9] Karl Gyllstrom and Marie-Francine Moens. Clash of the typings: Finding controversies and children’s topics within queries. In *Proceedings of the Advances in Information Retrieval - 33rd European Conference on IR Research.*, pp. 80–91, 2011.
- [10] Christoph Hölscher and Gerhard Strube. Web search behavior of internet experts and newbies. *Computer Networks*, Vol. 33, No. 1, pp. 337–346, 2000.
- [11] Moz Inc. 2020 Google search survey: How much do users trust their search results? <https://moz.com/blog/2020-google-search-survey>. 2022年10月7日閲覧.
- [12] Silvia Knobloch-Westerwick, Benjamin K. Johnson, and Axel Westerwick. Confirmation bias in online searches: Impacts of selective exposure before an election on political attitude strength and shifts. *Journal of Computer-Mediated Communication*, Vol. 20, No. 2, pp. 171–187, 2014.
- [13] Frances A. Pogacar, Amira Ghenai, Mark D. Smucker, and Charles L.A. Clarke. The positive and negative influence of search results on people’s decisions about the efficacy of medical treatments. In *Proceedings of the ACM SIGIR International Conference on Theory of Information Retrieval*, pp. 209–216, 2017.
- [14] 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*, pp. 1653–1662, 2019.
- [15] Christian Stab, Johannes Daxenberger, Chris Stahlhut, Tristan Miller, Benjamin Schiller, Christopher Tauchmann, Steffen Eger, and Iryna Gurevych. Argumentext: Searching for arguments in heterogeneous sources. In *Proceedings of the 2018 conference of the North American chapter of the association for computational linguistics*, pp. 21–25, 2018.
- [16] Yalin Sun, Yan Zhang, Jacek Gwizdka, and Ciaran B. Trace. Consumer evaluation of the quality of online health information: Systematic literature review of relevant criteria and indicators. *Journal of Medical Internet Research*, Vol. 21, No. 5, e12522, 2019.
- [17] 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, 2013.
- [18] Ryen W. White and Steven M. Drucker. Investigating behavioral variability in web search. In *Proceedings of the 16th International Conference on World Wide Web*, pp. 21–30, 2007.
- [19] Takehiro Yamamoto, Yusuke Yamamoto, and Sumio Fujita. Exploring people’s attitudes and behaviors toward careful information seeking in web search. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*, pp. 963–972, 2018.
- [20] 浜島聡一郎, 山本岳洋, 山本祐輔, 大島裕明. 健康情報検索における信憑性判断と意見の形成に関する調査. 第14回データ工学と情報マネジメントに関するフォーラム, A34–5, 2022.