

Web アンケートの形式と設問文の口調が 自由記述設問の回答に及ぼす影響

加藤 純世¹ 三山 貴也¹ 畑中 健壱¹ 中村 聡史¹

概要：Web アンケートは手軽に多くのデータを収集できる一方で、低品質な回答が混入しやすい。特に自由記述設問では意味のない文字列を入力するといった不適切な回答が生じやすい。こうした自由記述回答の質には、アンケートの形式や設問文の口調が影響を及ぼす可能性がある。具体的には、SNS の普及によりチャットや口語表現が身近になったため、チャット形式と口語的な口調の設問文を組み合わせることで回答者の心理的ハードルを下げ、より詳細な回答を引き出せることが期待される。一方、形式と口調の整合性が低い場合には回答者が設問に違和感を抱き、回答の質を損なう懸念もある。そこで本研究では、Web アンケートの形式および設問文の口調が自由記述設問の回答に及ぼす影響を調査する。「形式」にはチャット形式、標準形式という2つの水準、「口調」には口語、普通という2つの水準を設け、これらを組み合わせた4条件で実験を行い、自由記述設問の回答を分析した。その結果、回答トピック数については有意な差はみられなかった一方、回答時間についてはチャット形式において短くなることが示唆された。

1. はじめに

Web 上で実施するアンケートは、短期間かつ手軽に多くのデータを集めることができるというメリットがある。Nayak ら [1] は、Web アンケートは作成が容易であり、収集したデータの可視化も簡易であるとしている。このようなメリットから、Web アンケートでは多様な形式の設問を設計することが可能であり、多種多様な回答を集めることができると考えられる。そのため、大規模な基礎データの収集を目的として、学術研究や市場調査といった幅広い分野で活用されている。

しかし、Web アンケートでは不真面目な回答を行う回答者が多く存在するという問題がある [2]。特に自由記述設問では、回答者自身にとって回答可能な設問内容にもかかわらず、自由記述欄に「特になし」と記入したり、意味のない適当な文字列を入力したりするなど、不適切な回答が生じやすい [3]。Reja ら [4] は、Web アンケートの選択設問と自由記述設問の回答の質を比較をしたところ、自由記述設問の方が選択設問より欠損データが多く存在することを明らかにしている。このような不適切な回答は、得られるデータの信頼性を損ない、分析に用いるサンプル数を減少させる可能性がある。そのため、Web アンケートでは不適切な回答を抑制するための対策が求められる。

こうした背景を踏まえ、我々はこれまで自由記述設問における回答行動に影響を与える可能性のあるアンケート設計上の要因に着目し、多くの研究を行ってきた。具体的には、アンケート内における自由記述設問の提示位置やテキストボックスの表示サイズといった要素 [5]、PC やスマートフォンといった回答デバイスの違い [5, 6] に着目し、実験や分析を実施してきた。その結果、これらの要素が不真面目回答率やアンケートの途中での離脱率といった回答行動に影響を及ぼす可能性が示唆された。そしてこの結果から、Web アンケートには回答行動に影響を与える可能性がある多様な要素が存在していることが明らかになった。

ここで、回答行動に影響を与える要素として、アンケート形式と設問文の口調の2つが考えられる。アンケート形式については、従来の標準的な Web フォームに加え、対話的なやり取りを想起させるチャット形式のアンケートが用いられるようになっており、回答者の回答行動に影響を及ぼす可能性がある。また、設問文の口調についても、標準的で堅い表現と口語的な表現とでは回答者の回答のしやすさが異なると考えられる。

この点に関して、Kim ら [7] は、アンケート形式（標準/チャット）と設問文の口調（普通/口語）を要因とした2×2の実験を行い、選択設問のみからなるアンケートにおいて、チャット形式の回答者は標準形式の回答者よりも選択肢を分化して回答する傾向があり、さらに「チャット×口語」条件でその傾向が顕著になったことを報告している。

¹ 明治大学
Meiji University

この結果はアンケート形式および設問文の口調が回答行動に影響を与える可能性を示唆しているが、選択設問を対象としたものであり、自由記述設問における回答行動については十分に検討されていない。

そこで本研究では、アンケートの形式および設問文における口調に着目し、これらが自由記述設問の回答に及ぼす影響を調査する。具体的には、自由記述設問における回答トピック数を指標として、アンケートの形式（チャットと標準）と口調（口語と普通）が回答に及ぼす影響を分析する。

2. 関連研究

2.1 Web アンケートに関する研究

Web アンケートでの視覚的操作による影響については、多くの調査が行われている。Toepoel ら [8] は、1 画面に表示する選択設問数について検討を行い、1 画面当たりの設問数が多いほど無回答の数が増え、回答時間が短くなることを明らかにしている。Mavletova ら [9] は、17 問の設問が 2 ページで表示されるグループと、17 ページで表示されるグループでアンケートを実施したところ、2 ページのグループの方が回答時間が短く、アンケートの主観的評価を高めることを示している。

アンケートにおける設問の文章や事前に提示される文章に関する研究も多く行われている。Lenzner [10] は、理解しにくい質問を受けた人は、理解しやすい質問を受けた人より回答時間や回答の一貫性を指標とした回答の質が低下することを明らかにしている。永井ら [11] は、設問の前に付与する短い条件文が回答内容を変化させることを明らかにし、設問文の言い回しが回答者の判断に影響する可能性を指摘している。また Chaudhary ら [12] は、設問での重要性に関する記述とテキストボックスサイズが回答率に及ぼす影響を調査し、テキストボックスサイズに関係なく重要性の記述があることにより、Web アンケートでの回答文字数が有意に多いことを示している。

これらの研究は、Web アンケートにおいて、設問の提示方法や提示される文章が回答文字数や回答時間に影響を及ぼすことを示している。本研究は、アンケート形式という設問の提示方法、設問文の口調という提示される文章の違いが自由記述回答にどのような影響を及ぼすかを検証するものである。

2.2 アンケート形式に関する研究

アンケート調査における形式に関して、これまで様々な研究が行われている。医療分野における te Pas ら [13] の研究では、患者が標準形式のアンケートよりもチャットボットを用いたチャット形式のアンケートを好み、回答時間が同程度にもかかわらず、チャット形式の方が「より早く終わった」と知覚されることが示されている。また、チャッ

ト形式のアンケートは回答データの質を高め得ることが示唆されている。Xiao ら [14] は、自己紹介に関する自由記述設問を用いた比較実験を行い、チャットボットを用いたチャット形式の方が標準形式よりも回答量が多く、内容も充実した自己開示が得られたことを報告している。

これらの研究は、チャット形式というアンケート形式が回答量や回答内容に影響を及ぼすことを示している。本研究では、これまで十分に検討されていないアンケート形式と設問文の口調の組み合わせが、自由記述回答に及ぼす影響を調査するものである。

2.3 口調に関する研究

アンケート調査における口調に関しても、これまで様々な研究が行われている。Okoso ら [15] は、AI が提示する文章のカジュアル、フォーマル、ユーモラスといった複数の文体がユーザの意思決定に影響を与える可能性を示している。また、調査協力の場面でカジュアルな言葉遣いを用いると、回答者の心理的ハードルを下げる効果があることも報告されている。Song ら [16] は、中国の高校生を対象としてパネル調査を実施する中で、E メールやチャットで若者の言葉（インフォーマルな表現）を交えたところ、質問への参加姿勢が良好に保たれることを明らかにしている。また、Henson ら [17] は面接調査において、専門的で硬い口調より、親しみのある柔らかい口調の方が回答者が詳細な情報を開示しやすくなることを示している。

これらの研究は、アンケート調査における口調が回答者の心理状態や回答量に影響し、より詳細な回答を促すことを示している。本研究はこうした口調の効果がアンケート調査の自由記述回答にどのように影響するかを再検討し、さらにアンケート形式との交互作用を検証するものである。

3. 実験

3.1 実験概要

本研究では、アンケート形式と設問文の口調が自由記述回答に与える影響を調査するため、参加者間要因計画による 2 要因 2 水準のデザインを用いてアンケート調査を実施した。水準としてアンケート形式にはチャット形式、標準形式、設問文の口調には口語、普通を設定した。参加者には、これらを組み合わせた 4 条件のうち 1 条件に無作為に割り当てられアンケートに回答してもらった。本研究では以下の 2 つの仮説を設けた。

- **H1**：アンケート形式がチャットでかつ口調が口語であると、他の 3 条件に比べ回答トピック数が増える
- **H2**：アンケート形式が普通でかつ口調が口語であると、他の 3 条件に比べ回答トピック数が減る

3.2 アンケート構成

本研究のアンケートは、実験アンケートと事後アンケートの2つで構成した。実験アンケートで用いるテーマは「漫画」とした。テーマ設定および設問内容の作成に当たっては、Web アンケートにおける自由記述の特性を検討した Yamazaki ら [6] の研究を参考にした。

表1に示すように、実験アンケートは選択設問5問と自由記述設問2問の計7問で構成される。口語条件における設問文は、普通条件の設問文を著者らが口語的表現に書き換えたものである。

表2に、事後アンケートの設問内容の一覧を示す。これはユーザ属性を取得するためのアンケートで、ユーザの年齢、日常生活でチャットのやり取りをする頻度、その際の口調を問う選択設問3問で構成されている。なお、事後アンケートの設問文は、すべての実験条件で統一して普通の文体で提示した。

3.3 実験参加者

実験参加者は、アカウントの信頼性と完了率が高い [18] Yahoo!クラウドソーシング*¹ (YCS) を通じて募集した。YCS は、国内で広く利用されているクラウドソーシングサービスの一つで、高い完了率とアカウントの信頼性から、その有用性が示されている。実験を完了した参加者に62ポイントのPayPayポイントを報酬として支払った。また、Yamazaki ら [6] の研究においてPCとスマートフォンといった回答デバイスによっても回答行動に差がみられることが明らかとなっており、それを考慮して回答デバイスはスマートフォンに限定した。実験参加者は、男性100名、女性100名の計200名を募集した。

3.4 実験システム

本研究では、アンケート形式を要因として扱うため、チャット形式と標準形式の2種類のUIを実装した。

チャット形式のUIは図1に示すように、LINEのようなメッセージアプリに類似したものとなっている。選択設問では、設問文の直下に選択肢ボタンを表示する設計とし、チャットの流れに沿って自然に回答できるようにした(図1(a))。また、自由記述設問では、設問文は画面左側に吹き出しとして提示され、参加者の回答は画面下部に配置された入力欄から回答を送信することで、画面右側へ追加されるようにした(図1(b))。

標準形式のUIは図2に示すように、シンプルなものである。選択設問では設問文、選択肢を上から順に表示し、選択肢をタップすると次の設問の画面に遷移するようにした(図2(a))。自由記述設問では設問文、テキストボックス、送信ボタンを上から順に表示し、回答入力後に送信ボタン

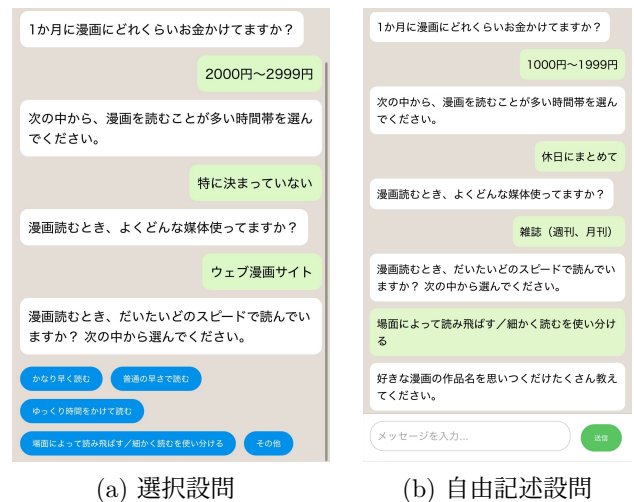


図1 チャット形式のアンケート画面

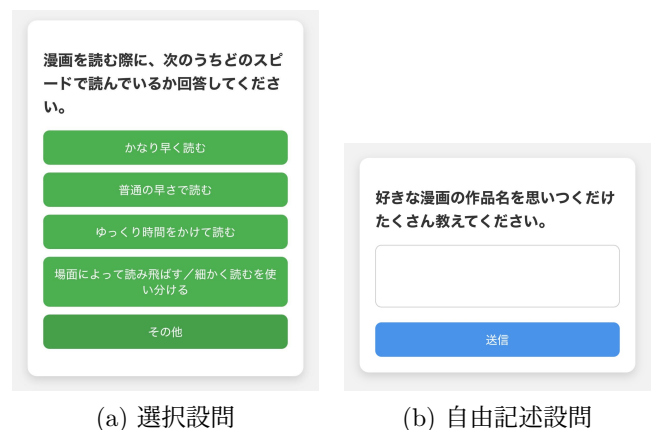


図2 標準形式のアンケート画面

を押すことで次の画面に遷移するようにした(図2(b))。

なお、事後アンケートでは、実験条件に関係なく標準形式のUIに統一した。本研究で使用した実験システムはすべてPHPとJavaScriptを用いて実装した。

3.5 実験手順

実験参加者は、YCS上に掲載されたタスク説明を確認した後、タスク画面に表示されるリンクから実験ページへアクセスしてもらった。参加者がアンケート開始画面に到達した時点で、実験条件の4条件のうち1条件をランダムに割り当てた。開始画面ではアンケート概要が提示され、参加者は「アンケートを開始する」ボタンを押すことで回答画面に遷移する。遷移後は、割り当てられた条件に応じて表示されるUIに従ってアンケートに回答し、7問の実験アンケートすべてに回答することで実験アンケートが完了するようにした。

実験アンケートが終了すると、事後アンケートへ進み、参加者がすべての設問の回答を終えると、報酬を受け取るためのコードとユーザ固有のIDが表示され、そのコードをYCSのタスク提出欄に入力することで報酬を受け取

*¹ <https://crowdsourcing.yahoo.co.jp>

表 1 実験アンケートの設問一覧

設問番号	回答形式	普通の設問文	口語の設問文
1	選択設問	漫画を一か月にどれくらい読むか回答してください。最近漫画を読んでいない人は過去に読んでいた経験で回答してください。	1 か月にどれくらい漫画読んでますか？もし最近読んでないなら、前に読んでたころの経験で答えてください。
2	選択設問	一か月あたりの漫画への出費額を教えてください。	1 か月で漫画にどれくらいお金かけてますか？
3	選択設問	次の中で、漫画を読む時間帯として多いものを選択してください。	次の中から、漫画を読むことが多い時間帯を選んでください。
4	選択設問	漫画を読むときの主な媒体について教えてください。	漫画読むとき、よくどんな媒体使ってますか？
5	選択設問	漫画を読む際に、次のうちのどのスピードで読んでいるか回答してください。	漫画読むとき、だいたいどのスピードで読んでいますか？次の中から選んでください。
6	自由記述	あなたが好きな漫画の作品名を、できるだけたくさん回答してください。	好きな漫画の作品名を思いつづけてたくさん教えてください。
7	自由記述	普段漫画を読む人は、漫画を読んでいて楽しいと感じる瞬間について教えてください。普段漫画を読まない人は、その理由や読まなくなったきっかけについて回答してください。	普段漫画読む人は、どんなときに「楽しい！」って感じますか？あんまり読まない人は、なんで読まないのかとか、読まなくなったきっかけについて教えてください。

表 2 事後アンケートの設問一覧

設問番号	回答形式	設問文	選択肢
1	選択設問	あなたの年齢が含まれるものを選択してください。	10-19/20-29/30-39/40-49/50-59/60-69/70-
2	選択設問	普段どれだけチャットのやり取りを行っていますか？	ほぼ毎日やり取りしている/週に数回程度やり取りしている/月に数回程度やり取りしている/ほとんどやり取りしない
3	選択設問	チャットにおいて普段どんな口調でやり取りをしていますか？	丁寧でかしこまった口調/くだけたカジュアルな口調/特に気にしない

ることができる仕組みとした。

4. 分析方法

4.1 自由記述回答の評価方法

表 1 に示すように自由記述設問は設問 6 と設問 7 の 2 問で構成されている。これらの回答の質を分析するため、本研究では回答内容を回答文字数ではなく回答トピック数に基づいて評価することにした。これは本研究で用いるチャット形式や口語といった要素が、それらに慣れ親しんでいる回答者の回答を冗長に長くする可能性を考慮したためである。

設問 6 は「あなたが好きな漫画の作品名を、できるだけたくさん回答してください」という、複数の語句（作品名）を列挙する形式の設問である。そのため、回答中に挙げられた作品数を回答トピック数として定義した。

一方、設問 7 は漫画を楽しんでいる瞬間、または読まない理由について文章で回答する形式の設問である。この設問 7 に対しては、回答文中に含まれる意味的に独立した内容の数を回答トピック数として定義した。この定義に関しては、Kleinkorres ら [19] による各回答を内容的に独立したアイデア単位として扱い、質的な評価を行う手法を参考にした。

回答トピック数の算出に当たっては、2 名の評価者によってすべての自由記述回答を評価した。評価者には、いずれの設問回答がどの実験条件に属するかがわからないように配慮し、さらに設問ごとに回答の提示順をランダム化した。

評価者 2 名が算出したトピック数の一致度は級内相関係数を用いて算出し、一致しなかった設問に対しては、著者らが回答トピック数を判断した。

4.2 除外データの基準

本研究においてはアンケートを途中で離脱した回答者のデータ、および設問に対して不適切な回答をした回答者のデータを分析から除外した。これらのデータは、すべての設問に対する回答が得られておらず、あるいは設問の意図を反映していないため、回答データを適切に評価することが困難であるうえ、分析に用いるデータの信頼性を歪めるためである。

該当するデータを除外した結果、アンケートの途中で離脱した回答者 8 名、設問に対して不適切な回答をした回答者 47 名の計 55 名が除外対象となり、分析対象は 145 名となった。また回答トピック数の評価に関して、評価者間の級内相関係数を算出した結果、設問 6 は 0.988、設問 7 は 0.719 となり、十分に信頼性があることが示された。

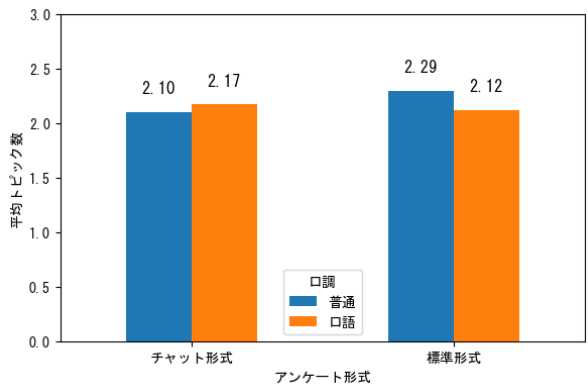


図 3 条件ごとの回答トピック数の平均

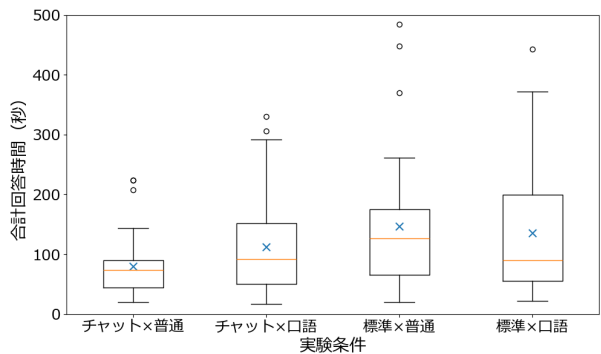


図 4 条件ごとの回答時間の分布

5. 結果

5.1 仮説に関する結果

図 3 に、各条件における自由記述回答の平均トピック数を示す。結果として、いずれの条件においても大きな差はみられず、全体としてほぼ同じ数値となった。また二元配置分散分析の結果、アンケート形式と設問文の口調のいずれの主効果も有意ではなく、両者の交互作用も見られなかった。この結果から、アンケート形式と口調の違いが回答トピック数に影響を与えているとは言えず、H1 および H2 は支持されなかった。そのため、以下では探索的分析の結果を報告する。

5.2 回答時間に関する結果

図 4 に、各条件における自由記述設問の合計回答時間の分布を示す。「チャット×普通」条件において、回答時間の平均値と中央値が最も低く、また分布の広がりも最も小さかったことから、多くの回答者が安定して短時間で回答していたことが観察された。一方、「標準×普通」条件では、平均値と中央値が最も高く、回答時間が長くなる傾向がみられた。さらに二元配置分散分析の結果、アンケート形式の主効果が有意であった ($p < 0.05$)。つまり、チャット形式は標準形式と比べて回答時間が有意に短かった。なお、口調の主効果および両者の交互作用は認められなかった。

表 3 トピック数ごとの回答者数

	1	2	3	4	5	6	7 以上
設問 6	42	24	29	13	14	6	7
設問 7	114	20	1	0	0	0	0

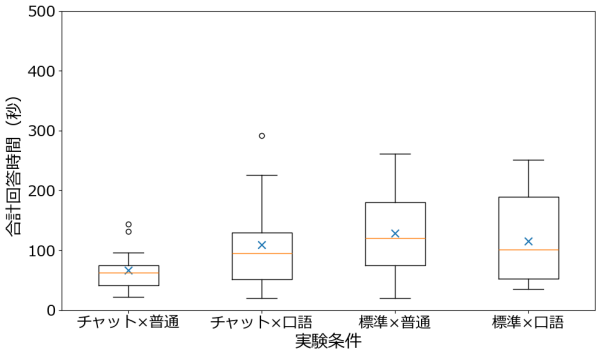


図 5 女性の回答時間の分布

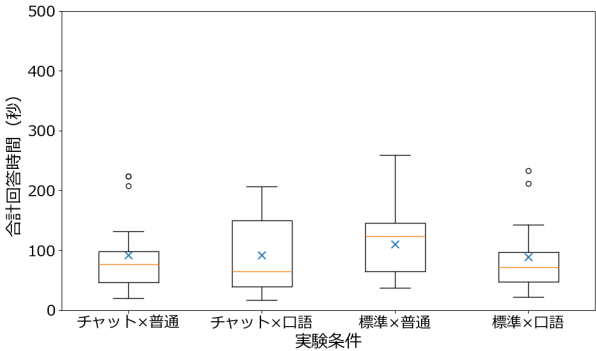


図 6 男性の回答時間の分布

5.3 ユーザ属性ごとの結果

本節では、ユーザ属性ごとの傾向を探索的に分析する。ここで、属性を加えて条件をさらに細分化することで 1 条件当たりのサンプル数が減少し、個々のユーザの回答行動が結果に与える影響が大きくなる。特に自由記述の回答時間においては、ごく一部の極端に長い回答時間が平均値を大きく歪める可能性がある。そこで、自由記述の回答時間が 300 秒を超えるユーザを外れ値として除外した。300 秒という基準は、自由記述設問の内容を踏まえ、通常の回答行動としては著しく長いと著者が判断し、設定したものである。基準値を超えたデータを除外した結果、10 名が除外対象となり、分析対象は 135 名となった。

図 5 に女性参加者の自由記述設問での合計回答時間の分布を示す。全体としてチャット形式よりも標準形式の方が回答時間が長く、分布のばらつきも大きい傾向がみられた。特に「標準×普通」条件では中央値および平均値が最も高く、四分位範囲も 2 番目に広いことから、個人差が顕著であった。一方、「チャット×普通」条件では中央値および平均値が最も低く、「標準×普通」に比べ分布も比較的安定しており、回答時間が短くなる傾向が確認された。また二元

表 4 実験条件ごとのトピック数

実験条件	設問 6	設問 7
チャット×普通	2.82	1.19
チャット×口語	3.16	1.08
標準×普通	2.91	1.26
標準×口語	2.78	1.24

配置分散分析の結果、アンケート形式の主効果が有意であり ($p < 0.05$)、女性参加者の回答時間はチャット形式の条件で回答時間が短くなる可能性が示された。

図 6 に男性参加者の自由記述設問での合計回答時間の分布を示す。全体として女性ほど条件間の差は小さくなく、各条件の平均値は比較的近い値を示した。「標準×普通」条件はやや回答時間が長いものの、他条件との明確な差は見られなかった。また二元配置分散分析の結果、アンケート形式、口調、およびその交互作用のいずれについても有意差は認められず、男性参加者の回答時間はアンケート形式や設問文の口調による影響を受けにくい可能性が示された。

以上の結果から、回答時間に関する影響は性別によって異なる可能性が示唆された。女性はチャット形式における回答時間が標準形式に比べ有意に短かったのに対し、男性では形式・口調のいずれの違いによっても回答時間に有意差は認められなかった。

図 7 に事後アンケート設問 3 で選択された選択肢ごとに、自由記述の平均文字数を実験条件別に整理した結果を示す。「標準×普通」条件を除き、「丁寧でかしこまった口調」を選択した参加者は、他の選択肢を選んだ参加者よりも平均回答文字数が多い傾向がみられた。一方、「くだけたカジュアルな口調」と回答した、日常的に口語的な表現を用いてチャットのやり取りをしている参加者は、普通条件で提示された設問文の方が文字数が多くなる傾向がみられた。またこの傾向は、標準形式において特に顕著であった。「丁寧でかしこまった口調」、「特に気にしない」を選択した参加者については、条件間で回答文字数に大きな差はみられなかった。

6. 考察

6.1 回答トピック数について

本研究は、アンケート形式と設問文の口調が自由記述設問の回答に及ぼす影響の調査という主目的のもと、H1「アンケート形式がチャットでかつ口調が口語であると、他の 3 条件に比べ回答トピック数が増える」、H2「アンケート形式が普通でかつ口調が口語であると、他の 3 条件に比べ回答トピック数が減る」を検証するものであった。

回答トピック数の分析においては、いずれの条件間でも有意差は認められず、明確な傾向も確認されなかった。表 3 は、自由記述設問における回答トピック数ごとの回答者数である。設問 6 では、回答によってトピック数にばらつ

きがあることが確認された。設問 7 では、100 名以上の参加者がトピック数 1 と評価されていた。この結果から、回答トピック数の分布は設問によって大きく異なっていることが示された。設問 6 は、語句を列挙する設問であり、挙げた語句の数がそのままトピック数となるため、トピック数のばらつきが大きくなったと考えられる。設問 7 は自身の経験を文章で記述する設問であり、意味的に独立した内容の数がトピック数となることから、トピック数のばらつきが小さくなったと考えられる。

このように、設問間でトピック数の分布に違いがみられたことから、設問内容の違いが実験結果に影響していた可能性を確認するため、設問別に実験条件ごとの回答トピック数の平均値を整理したものを表 4 に示す。その結果、いずれの設問においても、実験条件間で回答トピック数の平均値に大きな差はみられなかった。このことから、本研究で用いた自由記述設問が、アンケート形式や設問文の口調の違いによって回答トピック数に差が生じにくいタスクであった可能性が考えられる。

また、別の原因として、実験システムの特性が影響している可能性がある。本研究で用いた実験システムでは、参加者が回答を送信しても「いいですね」「ありがとうございます」といったフィードバックがなく、回答送信と設問提示が機械的に繰り返される構造となっていた。そのため、UI はチャットアプリケーションに類似していても、実際には回答者がチャット形式や口語表現が持つ効果を十分に得られなかった可能性が考えられる。

6.2 回答時間について

自由記述の回答時間における分析の結果、アンケート形式の主効果が有意となり、標準形式よりもチャット形式の方が有意に回答時間が短いことが明らかになった。また、探索的分析として参加者を女性と男性に分けて分析したところ、女性ではアンケート形式の主効果が確認され、同様の結果が得られた。一方男性では、アンケート形式、口調、交互作用のいずれにおいても有意な効果は認められなかった。

探索的分析から男女間で回答時間に差が出た原因として、回答文字数が考えられる。図 8 および図 9 に、女性と男性に分けた平均回答文字数を示す。どちらの結果においても、チャット形式より標準形式の方が、また口語口調より普通口調の方が文字数が多い傾向がみられた。しかし、いずれの条件においても男女の差が小さく、明確な傾向は確認されなかった。

また女性については、回答時間の分析でチャット形式における回答時間が有意に短い結果が得られたが、チャット形式の回答文字数が極端に少ないわけではない。標準形式との差はわずかであり、回答文字数が回答時間に影響している可能性は低い。なお、回答時間が短い要因として、

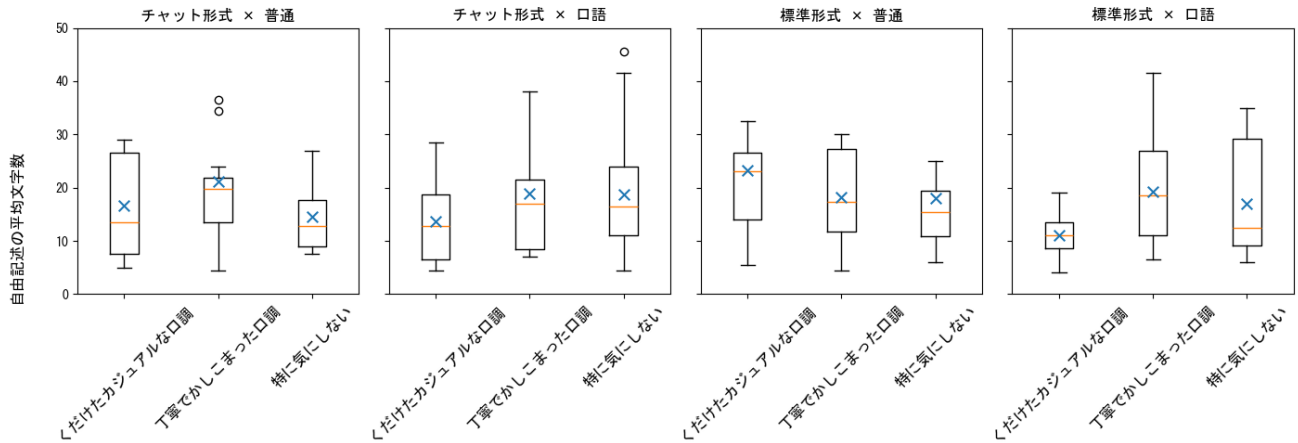


図 7 事後アンケート設問 3 の選択肢ごとの回答文字数

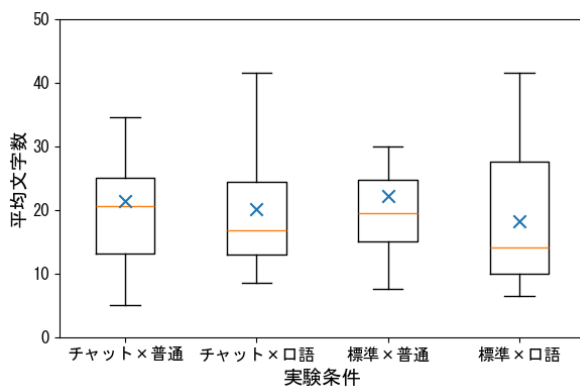


図 8 女性の回答文字数の平均

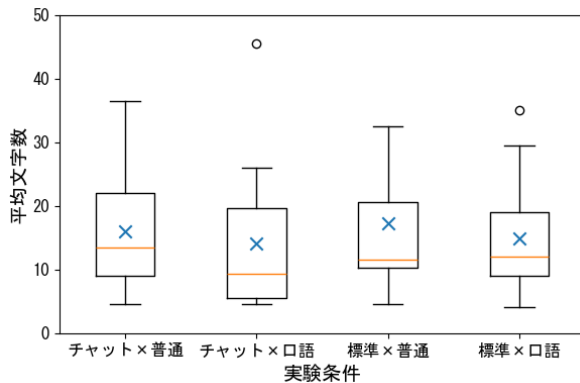


図 9 男性の回答文字数の平均

チャットをする頻度による慣れが考えられるが、事後アンケートより男女間でチャットをする頻度に大きな差はみられなかった。

ここで、チャット形式における女性参加者の回答時間が短くなった要因として、設問を提示するスピードが考えられる。本研究で用いたチャット形式のアンケートでは、参加者が回答を送信してから次の設問が表示されるまでの時間を 0.5 秒に設定しており、次の設問が早いテンポで提示されるようになっていた。女性の参加者が大半を占める研

究では、LINE におけるメッセージのやり取りを対象として、直前に受信したメッセージの返信時間と、次に自身が送信するメッセージの返信時間との間に正の相関があり、その原因は相手の早い返信に対して、自分の返信が遅いことへの罪悪感だと報告している [20]。以上の点を踏まえると、本研究においても設問が短い時間で提示されるチャット形式の構造が、女性参加者の回答時間に影響した可能性がある。つまりこの要因によって、女性参加者の回答を焦らせる効果が生じていた可能性がある。

今後は、設問文を提示するまでの時間を変化させるなどした場合に、回答時間や回答文字数がどう変化するかを検証する必要がある。

6.3 回答者が日常的に用いる口調との関連性

各条件の回答者を、事後アンケート設問 3 で選択された口調の違いごとに分類し回答文字数を分析した結果、「くだけたカジュアルな口調」を選択した回答者は、口語条件より普通条件の方が回答文字数が多い傾向がみられた。一方、「丁寧でかしこまった口調」、「特に気にしない」を選択した回答者には実験条件による明確な傾向はみられなかった。

この結果は、設問文の口調が常に回答量の増加につながるとは限らず、回答者が日常的に用いている口調との関係によって、その影響の現れ方が異なる可能性を示している。普段から口語的な表現でチャットのやり取りを行っている回答者にとっては、口語条件の設問文は日常的なやり取りに近いので、簡潔な表現で回答する傾向が強まった可能性がある。一方、普通条件では、より改まった表現で回答する必要があると認識され、回答文字数が増加した可能性があると考えられる。

ただし、この結果を解釈するにあたって、懸念点が存在する。それは、口語条件の設問文を著者らが作成した点である。本研究で用いた口語条件の設問文は、その妥当性を

十分に検討できていない。今後は、口語的な設問文の妥当性についても検証する必要がある。

7. おわりに

本研究は、Web アンケートにおけるアンケート形式と設問文の口調が自由記述設問の回答に及ぼす影響を調査した。その結果、回答トピック数に関しては、いずれの条件間においても有意な差はみられなかった。一方、探索的な分析の結果、回答時間についてはアンケート形式の主効果が確認され、チャット形式は標準形式と比較して有意に回答時間が短いことが明らかになった。また、性別ごとに分析を行ったところ、女性の参加者においてのみ回答時間が有意に短いことが確認され、男性の参加者では同様の結果は得られなかった。さらに、ユーザの口調属性に着目した分析からは、設問文の口調の効果が回答者の日常的に用いている口調によって異なる可能性が示唆された。

今後は、アンケートテーマや設問文の内容を変更した場合に、形式や口調による効果がどのように変化するかを検証する予定である。さらに、チャット形式における設問文の提示速度やフィードバックの有無による影響も検証する予定である。

参考文献

- [1] Nayak, M. S. D. P. and Narayan, K.: Strengths and Weakness of Online Surveys, *IOSR Journal of Humanities and Social Sciences (IOSR-JHSS)*, Vol. 24, No. 5, pp. 31–38 (2019).
- [2] Smyth, J. D., Dillman, D. A., Christian, L. M. and McBride, M.: Open-ended questions in web surveys: Can increasing the size of answer boxes and providing extra verbal instructions improve response quality?, *Public Opinion Quarterly*, Vol. 73, No. 2, pp. 325–337 (2009).
- [3] Holland, J. L. and Christian, L. M.: The influence of topic interest and interactive probing on responses to open-ended questions in web surveys, *Social Science Computer Review*, Vol. 27, No. 2, pp. 196–212 (2009).
- [4] Reja, U., Manfreda, K. L., Hlebec, V. and Vehovar, V.: Open-ended vs. close-ended questions in Web questionnaires, *Advances in methodology and statistics*, Vol. 19, No. 1, pp. 159–177 (2003).
- [5] Yamazaki, I., Hatanaka, K., Nakamura, S. and Komatsu, T.: The Effects of Order and Text Box Size of Open-ended Questions on Withdrawal Rate and the Length of Response, *35th Australian Conference on Human-Computer Interaction*, pp. 446–453 (2023).
- [6] Yamazaki, I., Hatanaka, K., Nakamura, S. and Komatsu, T.: The Effects of Order and Text Box Size of Open-ended Questions on Withdrawal Rate and the Length of Response, *35th Australian Conference on Human-Computer Interaction*, pp. 446–453 (2023).
- [7] Kim, S., Lee, J. and Gweon, G.: Comparing Data from Chatbot and Web Surveys: Effects of Platform and Conversational Style on Survey Response Quality, *Proceedings of the 2019 CHI Conference on Human Factors in Computing System*, No. 86, pp. 1–12 (2019).
- [8] Toepoel, V., Das, M. and Soest, A. V.: Design of Web Questionnaires: The Effects of the Number of Items per Screen, *Field Methods*, Vol. 21, No. 2, pp. 200–213 (2009).
- [9] Mavletova, A.: Data Quality in PC and Mobile Web Surveys, *Social Science Computer Review*, Vol. 31, No. 6, pp. 725–743 (2013).
- [10] Lenzner, T.: Effects of Survey Question Comprehensibility on Response Quality, *Field Methods*, Vol. 24, No. 4, pp. 409–428 (2012).
- [11] 永井大樹, 後藤隆: 調査における二種類の質問紙調査の回答の比較: 発育発達に関わる ライフスタイルの要因の文脈効果, 順序効果, 自由回答, 発育発達研究, Vol. 2009, No. 44, pp. 1–7 (2009).
- [12] Chaudhary, A. K. and Israel, G. D.: Influence of Importance Statements and Box Size on Response Rate and Response Quality of Open-Ended Questions in Web/Mail Mixed-Mode Surveys, *Journal of Rural Social Sciences*, Vol. 31, No. 3, pp. 140–159 (2016).
- [13] Pas, M. E. T., Rutten, W. G. M. M., Bouwman, R. A. and Buise, M. P.: User Experience of a Chatbot Questionnaire Versus a Regular Computer Questionnaire: Prospective Comparative Study, *JMIR Med Inform*, Vol. 8, No. 12 (2020).
- [14] Xiao, Z., Zhou, M. X., Liao, Q. V. and Mark, G.: Tell Me About Yourself: Using an AI-Powered Chatbot to Conduct Conversational Surveys with Open-ended Questions, *Proceedings of the 2025 CHI Conference on Human Factors in Computing System*, Vol. 27, No. 15, pp. 1–37 (2020).
- [15] Okoso, A., Yang, M. and Baba, Y.: Do Expressions Change Decisions? Exploring the Impact of AI's Explanation Tone on Decision-Making, *Proceedings of the 2025 CHI Conference on Human Factors in Computing System*, No. 824, pp. 1–22 (2025).
- [16] Song, J., Lai, G., Wong, O. and Feng, X.: Staying connected with ICT tools: tracking youth respondents in a Chinese context, *The Journal of Chinese Sociology*, Vol. 5, No. 1 (2018).
- [17] Henson, R., Cannell, C. F. and Lawson, S.: Effects of Interviewer Style on Quality of Reporting in a Survey Interview, *The Journal of Psychology*, Vol. 93, No. 2, pp. 221–227 (1976).
- [18] Seaborn, K. and Nakamura, S.: Quality and representativeness of research online with Yahoo! Crowdsourcing, *Frontiers in Psychology*, Vol. 16 (2025).
- [19] Kleinkorres, R., Forthmann, B. and Holling, H.: An Experimental Approach to Investigate the Involvement of Cognitive Load in Divergent Thinking, *Journal of Intelligence*, Vol. 9, No. 1 (2021).
- [20] 村上幸史: 返信は早い方が良いのか? —携帯メールやLINEにおける「互酬性仮説」の検証, 実験社会心理学研究, Vol. 62, No. 2, pp. 80–93 (2023).