

Web 上での調査における回答時間に着目した 不適切な慣れの基礎調査

高久拓海¹ 小松原達哉¹ 山崎郁未¹ 中村聡史¹

概要 : Web 上でのアンケート調査やユーザ実験, データセット構築など様々な用途でクラウドソーシングが活用されている。ここで, クラウドソーシング経由で, Web 上のシステムを用いた実験などの作業をしてもらう場合に問題となるものに, 作業中の慣れがある。作業中に生じる慣れの具体的な事例として, 作業に慣れてきたことでデータ登録を深く考えず行ってしまうことや, 単純作業であるがためにリズム感覚で回答を行ってしまうといったことが考えられる。そこで, 我々は作業中の慣れが実験者のデータ登録に及ぼす影響について研究を行う。そのための基礎調査として, これまでに実施してきた3つの選択実験を用いて作業中の回答時間に着目した分析を行うことで, 回答時間の傾向から慣れの影響を検出することの可否について調査を行った。分析の結果, リズム化グループは序盤から終盤にかけて回答時間が短くなり, また終盤では同じ選択位置を連続して選択するなど, 不適切な回答をすることが明らかになった。

キーワード : データセット, 慣れ, リズム化, 回答時間

1. はじめに

Web 上でのアンケート調査やユーザ実験, データセット構築など様々な用途でクラウドソーシングが活用されている。クラウドソーシングのサービスとしては, Yahoo!クラウドソーシング[1]や Amazon Mechanical Turk[2]などが知られており, 日々様々なタスクが膨大な登録ユーザにより実施されている。我々の研究室でもこれまで選択誘導[3]や選択肢におけるダークパターンの問題[4]などを調査するため, クラウドソーシングを活用して Web 上で実験を行っている。

Web 上でのアンケート調査やユーザ実験, データセット構築などは, 対面で実施するものに比べ, 安価で即座に膨大なユーザに依頼することが可能であるが, データの質という面で問題が生じている。例えば Web 上でのアンケートや作業の信頼度について, 三浦ら[5]は Web 上でのアンケート調査を実施し, ある会社では 83.8%の回答者が設問の内容に従わずに回答を行っていることを明らかにしている。また山崎ら[6]は, Web アンケートにおいて自由記述設問が後半にある場合や, Web アンケートが公開されてすぐの回答者は不真面目な回答をする傾向があることを明らかにしている。このように Web 上でのアンケートや作業は, 監督者に監視されず一人で作業を進めることが可能な環境であることが多いため, 質の悪いデータを登録する不真面目な協力者が存在することが考えられる。

先述した不真面目な協力者とは別に, 真面目に作業を行っていたにも関わらず, 意図せず何らかの影響を受けてしまい, アンケートへの回答や実験における行動, データセットの登録基準などの作業が変容してしまうことも考えられる。このような影響を受けた作業データもまた, アンケートや実験, データセットの質を下げる要因となる。

意図しない影響の一つとして, ある作業や動作を繰り返

し行っている際に生じる慣れがある。この慣れは, 心理学の概念の一つで馴化と呼ばれており, ある刺激が繰り返し提示されることによって, その刺激に対する反応が徐々に見られなくなる現象のことである[7]。馴化の特徴として, 刺激の提示頻度が多くなるほど馴化は早く進むことや, 刺激の強度が強くない場合だと馴化が早く進むこと, 同一の刺激が長時間与えられないと, その刺激に対する反応が自然に回復することなどが分かっている[8]。

こうした慣れは, アンケート調査やユーザ実験, データセット構築の参加者が作業を行うときにも起こると考えられ, 慣れによって登録されるデータに影響が生じる可能性がある。データ登録中に生じる慣れの具体的な例として, 提示される刺激を繰り返し見たことで次の刺激を先読みして登録作業を行うことや, ひたすらクリック作業を行うものであるため思考が間に合っていないのにデータ登録してしまうことなどが考えられる。このような慣れの影響をなくすため, 作業中に複数の異なる刺激を提示することで, 脱馴化を目的とする研究も行われている[9]。また, 慣れによるミスを防ぐための手法として, 指差し呼称を用いて注意力を向上させる研究も行われている[10]。しかし, 別の刺激の提示や別の行動を促すことは, 本来の登録作業に影響を及ぼす恐れもある。こうした慣れの問題を解決するには, 作業を行っている人に慣れによる影響が確認された場合のみ自覚させ, 改善させる必要があると考えられる。

この Web 上での調査や実験などにおける慣れの問題については, まだどういった特性があるのか十分に明らかになっていない。そこで本研究では, 慣れの影響の一つである作業を繰り返し行うことで設問への回答がリズム化してしまい, 回答が不適切になってしまうという問題に着目し, 基礎調査を行う。ここでは, 我々の研究室でこれまで実施してきた様々な実験のうち, 繰り返し選択を行う実験

¹ 明治大学
Meiji University

[4][11][12]に着目し、作業を行っている登録者が選択回答をリズム感覚で行った結果、回答が不適切なものになってしまっていると予想し、その分析を行う。具体的には、「回答作業中に慣れが生じてしまい、回答がリズム化した実験協力者は回答時間も短くなり、回答も不適切になる」という仮説を立て、実験データを多面的に分析することにより調査を行う。

2. 関連研究

2.1 データセットに関する研究

データセット内のデータの量や質に関する研究は様々行われている。Cai ら[13]は、データを分析および活用してその価値を保証するためにはデータセット内のデータの質を高く保つ必要がある点に着目した研究を行っている。彼らはデータの質が抱える課題を分析し、可用性、ユーザビリティ、信頼性、信用性、プレゼンテーションの質からなる、評価モデルの基礎となるフィードバック機構を備えた動的なビックデータ品質評価プロセスを定式化している。Chawla ら[14]は、データセット内のデータの偏りが機械学習やビジネス産業への利用に問題がある点に着目し、提案されている複数の解決策について、オーバーサンプリングや 1 クラス学習が有効であることを明らかにしている。Latch ら[15]は、データ数の少ないデータセットでは学習性能の低下や予測モデルの不確実性を引き起こすという点に着目し、この問題を解決するための研究を行っている。その結果、足りないデータを人工的に生成し補う人口データ生成法が対処法の一つとして有効であることを明らかにしている。松井ら[16]は、教師ありの機械学習を行う場合において公開されたデータセットが小さいという問題に着目し、十分な量の教師データを集めるため、フィジカルなゲームを楽しんでもらいながら、最終的に人間の動作をスケルトンデータとして収集することでデータセットの増築を行っている。このように、データセットの質を高く保つことを目的とした研究は多く行われている。

2.2 慣れに着目した研究

慣れに着目した研究は様々行われている。溝口ら[17]は、聴覚への刺激として騒音に対する慣れに着目した研究を行っている。具体的には、音の物理的な特性と慣れの関係を調べる手法として、音の物理量を制御した音刺激を用いて、比較的短い時間の実験で慣れを定量的に捉える手法を提案している。菊池ら[18]は、音声インタフェースの設計指針を明らかにするため、事前経験といった習熟の慣れが被験者の操作や評価に与える影響について評価実験を行った。その結果、これらの慣れを考慮する場合にはインタフェースの利用回数や利用間隔を考慮する必要があることを示している。高田ら[19]は、「視線入力とハンドジェスチャを併用したインタラクション」と「ハンドジェスチャによるイ

ンタラクション」において、慣れがどのような効果を及ぼすか調査を行っている。実験の結果、操作訓練と本番実験を行うことで各操作方法における操作時間や誤操作の回数が減り、一定の慣れの効果が得られることを明らかにしている。杉邑ら[20]はソフトウェアのユーザインタフェースに対するユーザの慣れに着目し、視線を利用することでこの慣れを定量的に測定する研究を行っている。実験の結果、ユーザインタフェースが慣れている場合とそうでない場合は、ある一定の範囲を見続ける注視時間に有意な差があることを明らかにしている。このように、慣れによって人の行動に変化が生じることがわかる。

2.3 繰り返し作業に関する研究

繰り返し行われる作業に着目した研究も多数行われている。Wood ら[21]は、作業と休息の時間に着目した研究を行っており、実験を通して作業と休息の開始時と終了時の疲労具合を予測するモデルを開発している。また、与えられた反復作業において疲労を最小化する作業と休息スケジュールの提案も行っている。Smith ら[22]は、軽い作業を繰り返し行う場合の作業量に着目し、作業の量が実験者に作業の完了を促すという仮説を立て、機械実験を行った。その結果、作業量が多いほど実験者の作業停止の回数が多い傾向であることを明らかにしている。辛島ら[23]は、単純繰り返し作業と音楽の関係性に着目し、作業前に音楽を聴取することでタスクのパフォーマンスが向上するかについて調査を行った。その結果、作業前に音楽を聴取することで作業時のパフォーマンスが向上する傾向があることを明らかにしている。このように、繰り返し作業における課題やその対処方法については研究が行われている。しかし、繰り返し作業中の慣れに着目した研究はこれまで行われていない。

3. 分析方針

3.1 対象となる分析データの特徴

本研究で述べる分析や調査がどのようなデータを対象としているかについて述べる。今回の分析で扱うデータは、クラウドソーシング経由で実施された Web 上での 3 つの異なる実験によって得られたものである。ここで扱う実験は、内容がすべて異なる設問に対し、実験協力者が一番適していると思う回答を選択するものである。また、作業中表示されるのは設問の内容と選択肢、選択肢を選ぶうえでの指示のみであり、選択肢を選ぶ作業はすべてクリック、またはタップで行えるものである。これらの作業では、クリック（またはタップ）を繰り返し行うものであり、その作業が単調であるため、リズム感覚で回答してしまう可能性がある。

3.2 分析のためのデータ分割

対象となるデータがリズム感覚で行われているかを分

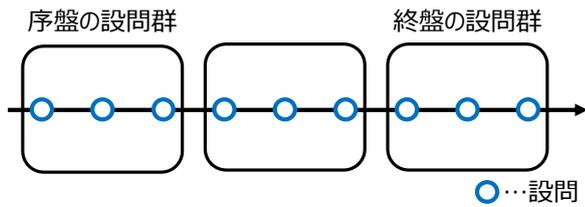


図1 序盤群と終盤群の例

析する際に、データをどのように扱うかについて述べる。

3.1 節で述べた通り、今回の分析で扱うデータは繰り返し作業によって登録されるデータである。そのため、作業の前半では影響を受けていないデータが、作業の後半では影響を受けているデータが登録されている可能性があると考えられる。

そこで本研究では、繰り返し作業において登録されたデータを設問番号順に群に分類する。ここでは、序盤の設問群（以降、序盤群と呼ぶ）と終盤の設問群（以降、終盤群と呼ぶ）に分け、その両者を比較することで回答者の分類を行う（図1）。

3.3 回答者分類の指標

選択にかかる時間がある区間においてリズム化しているかどうかを判断する材料として、その設問群内での選択時間の標準偏差を利用する。



図2 クリックによる選択操作の時間軸

ここで図2に示すように、各設問の回答時間を*i*番目の設問の選択時間を s_i 、分析対象の設問群を*m*番目から*n*番目 ($1 \leq m \leq n$)としたときの平均回答時間 $avg(m, n)$ と、回答の標準偏差 $\sigma(m, n)$ は次の式のように求められる。

$$avg(m, n) = \frac{\sum_{i=m}^n s_i}{n - m + 1}$$

$$\sigma(m, n) = \sqrt{\frac{\sum_{i=m}^n (s_i - avg(m, n))^2}{n - m + 1}}$$

設問の内容に独立性がある場合、真面目に回答しておりリズム化していなければ、設問ごとに回答時間が異なり、 $\sigma(m, n)$ の値は高くなる。一方、回答がリズム化されている場合は、設問の内容に関わらず回答が等間隔で行われるため、 $\sigma(m, n)$ の値が低くなる。この $\sigma(m, n)$ の値を用いて回答者分類を行う。

3.4 回答者分類

3.3 節で述べた序盤群と終盤群ごとに*m*と*n*の値を設定し、

$\sigma(m, n)$ の値を比較することで、実験者を以下の3つのグループに分類する。具体的には、序盤群を $m_{opening}$ から $n_{opening}$ まで、終盤群を $m_{closing}$ から $n_{closing}$ までとしたとき、ある閾値 α と β ただし($\alpha \leq \beta$)を用いて、下記の条件で分類する。

- **真面目グループ**: すべての設問について真剣に考えて回答した実験協力者のグループ。作業の序盤から終盤まで設問の内容に応じて回答を考えるため、回答時間が序盤から終盤までバラつくと考えられる。そこで、 $\sigma(m_{opening}, n_{opening}) \geq \beta$ かつ $\sigma(m_{closing}, n_{closing}) \geq \beta$ という条件で分類する。
- **リズム化グループ**: 作業の序盤は真面目に回答していたが、後半になるにつれてリズム化して回答した実験協力者のグループ。序盤は設問に応じて内容を考えるため回答時間にバラつきが見られるが、作業への慣れで後半部分での回答が無意識のうちにリズム化し、回答時間が等間隔になっていくと考えられる。そこで、 $\sigma(m_{opening}, n_{opening}) \geq \beta$ かつ $\sigma(m_{closing}, n_{closing}) < \alpha$ という条件で分類する。
- **不真面目グループ**: 序盤から作業を真剣に取り組む気がない実験協力者のグループ。報酬目当てでとにかく作業を早く終わらそうとするため、すべての設問に対して深い考えがなくクリック選択していった結果、回答時間が一定になると考えられる。そこで、 $\sigma(m_{opening}, n_{opening}) < \alpha$ かつ $\sigma(m_{closing}, n_{closing}) < \alpha$ という条件で分類する。

4. 分析対象

本研究で分析を行うデータセットについて、データ取得時の実験概要、実験の設問数、実験での選択肢、3.2 節で述べた設問群分け、3.4 節で述べた実験者のグループ分けのための、序盤群の($m_{opening}, n_{opening}$)と、終盤群の($m_{closing}, n_{closing}$)について説明を行う。

なお、これらの実験は Yahoo!クラウドソーシング[1]で依頼をしたものである。また、それぞれの実験では実験ページの最後に表示される ID をクラウドソーシング上の元ページで入力し、実験を終了することになっていたが、ID 入力を行わなかった実験協力者がいた。今回の分析は、そうした実験協力者も含めて分析を行う。

4.1 データセット A

データセット A は木下ら[4]が行った、選択肢の遅延表示が選択に及ぼす影響の調査実験の結果である。全部で 15 問の設問からなり、まずは設問を読んで、次へのボタンをクリックし、推移した先のページで 6 つの選択肢が表示される。実験協力者はその選択肢から対象を選ぶことを要求されるものである。実験の流れを図3に示す。

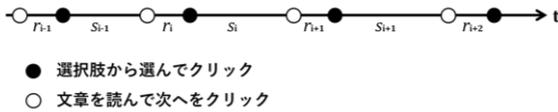


図3 データセット A の実験の流れ

ここで、設問が表示されてから選択肢を選ぶまでの流れは同じであるため、実験の後半に作業がリズム化してしまう可能性が考えられる。そこで実験データを、1~5 問目、6~10 問目、11~15 問目の 3 つに分割し、序盤群を 1~5 問目、終盤群を 11~15 問目とし、それぞれの $\sigma(m, n)$ の値で回答者を分類する。なお、本実験協力者は 2,037 名（男性 1,019 名、女性 1,018 名）であった。

4.2 データセット B

データセット B は植木ら[11]が行った、フォントが選択に及ぼす影響の調査を 3 択で行う実験の結果である。全部で 10 問の設問からなり、各設問のページで選択に関する説明文と 3 つの選択肢が表示されている。実験協力者が選択肢をクリックすると、次の設問と選択肢が提示されるため、順に選択だけを行っていくものである。実験の流れを図 4 に示す。

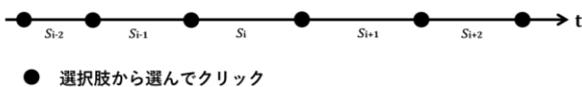


図4 データセット B の実験の流れ

ここで、設問が表示されてから選択肢を選ぶまでの流れは 10 問の設問で同じであるため、実験の後半に作業がリズム化してしまう可能性が考えられる。そこで実験データを、1~4 問目、5~6 問目、7~10 問目の 3 つに分割し、序盤群を 1~4 問目、終盤群を 7~10 問目とし、それぞれの $\sigma(m, n)$ の値で回答者を分類する。なお、本実験協力者は 1,963 名（男性 1,342 名、女性 621 名）であった。

4.3 データセット C

データセット C は関口ら[12]が行った、色が選択行動に及ぼす影響の調査実験の結果である。この実験は、全部で 20 問の設問からなり、4 問目、8 問目、12 問目、16 問目、20 問目では 3 つの選択肢が、それ以外では 2 択の選択肢が表示される。こちらはデータセット A と同じく、設問文を読んだ後、提示されているボタンをクリックし、その後に選択肢をクリックするという流れになっている（図 5）。

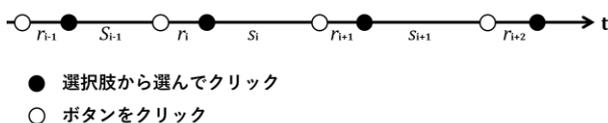


図5 データセット C の実験の流れ

ここで、設問が表示されてから選択肢を選ぶまでの流れは 20 問の設問で同じであること、また、2 択の設問と 3 択の設問の順番が固定であることから、実験の後半に作業がリズム化してしまう可能性が考えられるが、この実験ではシステムの回答方法がわかりにくいものであったため、1 問目は他に比べ選択時間がかなり長かった。そこで実験データを、1~4 問目、5~8 問目、9~12 問目、13~16 問目、17~20 問目の 5 つに分割したうち、1~4 問目を練習と捉え、序盤群を 5~8 問目、終盤群を 17~20 問目とし、それぞれの $\sigma(m, n)$ の値で回答者を分類する。なお、本実験協力者は 855 名（男性 439 名、女性 416 名）であった。

5. 分析結果

データセット A, B, C の分析結果について、各グループの分布人数、各設問群における各グループの平均回答時間、各選択肢の選択位置における連続回答の回数を述べる。ここで、データセット A と B については、すべての設問で選択肢の数が同じであったため、よりリズム感覚で選択を繰り返している可能性がある。そこで、各グループ各設問群でどれくらい同じ選択位置を連続で回答していたのかという点に着目して分析を行う。

5.1 各データセットにおけるグループの分布

本分析では、分類する際の α と β の値は、それぞれ $\alpha = 0.5$ 、 $\beta = 1.0$ とした。データセット A, B, C それぞれの真面目グループ、リズム化グループ、不真面目グループ、その他

表1 データセット A におけるグループごとの人数

グループ	人数 (人)
真面目	840
リズム化	78
不真面目	63
その他	1056

表2 データセット B におけるグループごとの人数

グループ	人数 (人)
真面目	847
リズム化	128
不真面目	73
その他	915

表3 データセット C におけるグループごとの人数

グループ	人数 (人)
真面目	124
リズム化	66
不真面目	196
その他	469

グループ(いずれのグループにも含まれないもの)が表 1, 2, 3 の通りとなった. この結果より, いずれの実験においてもリズム化グループ, 不真面目グループが一定数以上いることがわかる.

5.2 各設問群における平均回答時間

データセット A における, 各グループにおける設問群の平均回答時間の推移を図 6 に示す. なお, 真面目グループにおいて回答時間が 20 秒を超える外れ値が観測されたため, 一度でも 20 秒を超える回答をした回答者のデータは, 除外したうえで分析を行った. この図より, グループごとに回答時間を比較すると, 真面目, リズム化, 不真面目グループの順に長いことがわかる. また, 真面目グループと不真面目グループは, 序盤群, 中盤群, 終盤群の平均回答時間はほぼ変わらない一方で, リズム化グループでは, 序盤群から終盤群にかけて回答時間が徐々に減少していることがわかる.

データセット B における, 分類を行った各グループにおける序盤 4 問, 中盤 2 問, 終盤 4 問の平均回答時間の推移を図 7 に示す. この図より, データセット A と同じく真面目, リズム化, 不真面目の順に回答時間が長くなった. また, 真面目グループと不真面目グループは平均回答時間に変化がないが, リズム化グループでは序盤群から終盤群にかけて平均回答時間が減少していることがわかる.

データセット C における, 分類を行った各グループにおける 5~8 問目, 9~12 問目, 13~16 問目, 17~20 問目の平均回答時間の推移を図 8 に示す. この図より, 真面目グループがすべての設問群において最も回答時間が長い結果となったことがわかる. また, 真面目グループと不真面目グループは平均回答時間にあまり変化がないが, リズム化グループでは序盤群から終盤群にかけて平均回答時間が徐々に減少していることがわかる.

ここで, データセット C は 2 択と 3 択とが入れ替わるものであったため, その各グループについて 1~20 問目における設問ごとの回答時間の平均を図 9 に示す. ここで, 設問 1 について, 真面目グループとリズム化グループの値が高いことがわかるが, これは先述の通り, システムの回答方法がわかりにくいものであったためである. 次に, 設問番号が 4 で割り切れる 3 択のものに着目すると, 真面目グループではその前の 2 択の設問に比べ選択時間が伸びていることがわかる. また, リズム化グループでも 3 択設問では直前の設問の回答時間より長くなる傾向が見られたが, 20 問目ではその増加が見られなかった. 一方不真面目グループでは, 回答時間の大きな増加や減少は見られなかった.

以上の結果より, いずれのデータセットにおいても, 回答がリズム化した実験協力者は, 回答時間が短くなることが確認された.

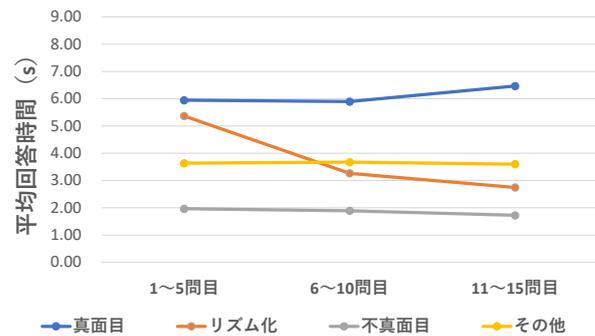


図 6 データセット A の平均回答時間

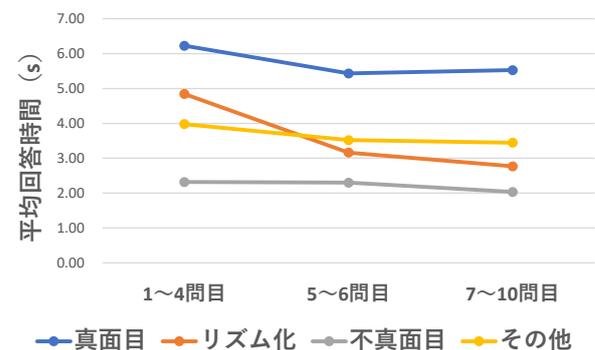


図 7 データセット B の平均回答時間

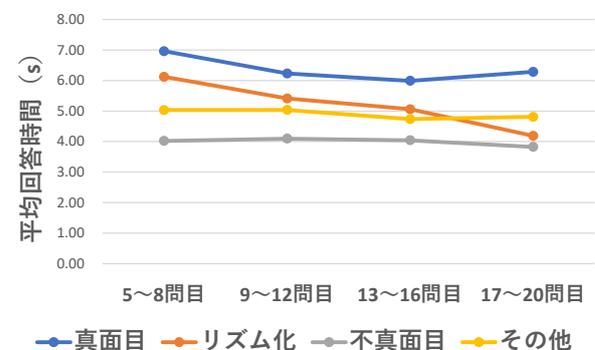


図 8 データセット C の平均回答時間

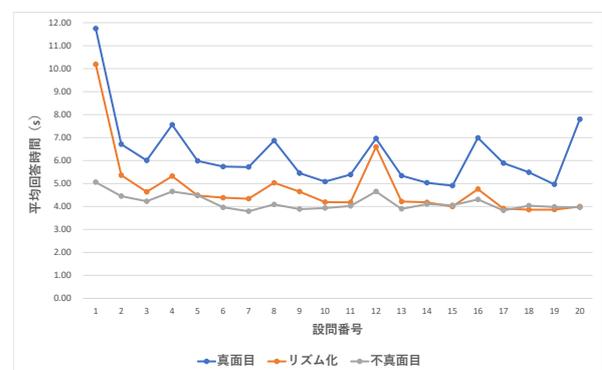


図 9 データセット C の設問ごとの平均回答時間

5.3 選択位置による連続回答

分類されたグループによって選択傾向に影響があるかどうかを調査するため、データセット A, B, C における選択傾向を分析する。

データセット A について、各グループで左右位置が同じ選択位置（本来 6 択だが、左・中・右の 3 択として扱う）を 3 回以上連続して回答した人数の割合を、グループおよび群ごとに分析した結果を図 10 に示す。この結果より、真面目グループはすべての設問群で 3 回以上連続して回答した人数の割合が一番低く、不真面目グループはすべての設問群で値が一番高いことがわかる。また、リズム化グループでは、序盤群に比べ終盤群が高い値となり、序盤より終盤の方が、選択が不適切になっていると考えられる。

データセット B について、各グループで同じ選択位置を 3 回以上連続して回答した人数の割合を、グループおよび群ごとに分析した結果を図 11 に示す。この結果より、序盤群では真面目グループが、終盤群では不真面目グループが 3 連続以上続けて回答した割合が一番高いことがわかる。一方、リズム化グループでは、データセット A と同じく序盤群に比べ終盤群が高い値となり、序盤より終盤の方が、選択が不適切になっていると考えられる。

データセット C について、2 択の設問部分にのみ着目し、選択傾向に影響があるかを序盤群の 5～7 問目、終盤群の 17～19 問目についてグループごとに分析した結果を図 12 に示す。この結果より、リズム化グループと不真面目グループは終盤群で値が高くなる傾向が見られた。一方、真面目グループでは、序盤群と終盤群で値に違いがない結果となった。

以上の結果より、いずれのデータセットにおいても、回答がリズム化した実験協力者は、回答も不適切になることが確認された。

6. 考察

5 章で得られた分析結果をもとに、各分析における真面目、リズム化、不真面目グループの傾向と、今後の課題について考察を行う。

5.2 節の図 3, 図 4, 図 5 より、データセット A, B, C のすべての結果で、回答時間は各設問群において、真面目、リズム化、不真面目グループの順に長い結果となった。そこで、データセットごとに平均回答時間の推移について考察を行う。

まず 5.2 節の結果について、真面目グループは、実験を通して設問に対する回答時間が長い傾向が見られた。これらのことから、選択のリズム化が起こることなく、実験が進んでも選択肢を悩んでいることが考えられる。一方、不真面目グループは実験を通して回答時間が短い傾向があり、序盤からマウスを連打するような選択を行っていることが

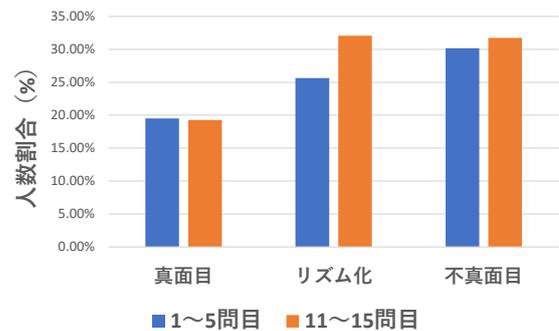


図 10 データセット A の選択肢の連続選択人数

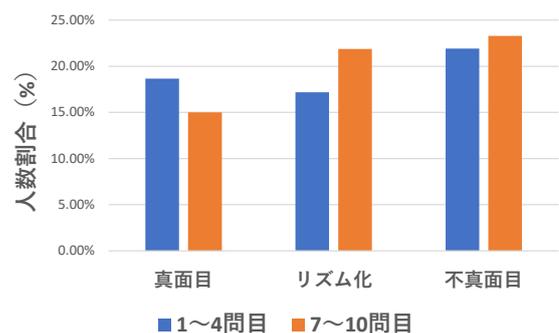


図 11 データセット B の選択肢の連続選択人数

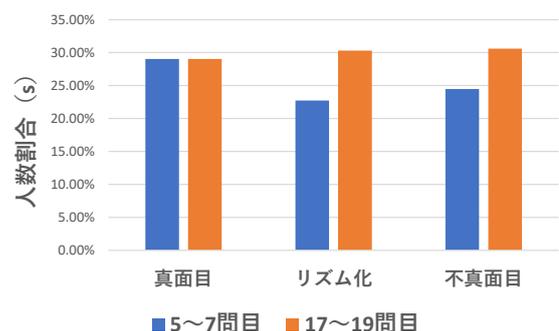


図 12 データセット C の選択肢の連続選択人数

予想される。またリズム化グループは、序盤群において回答時間が長いものの、終盤群では回答時間が短いことがわかった。これは、序盤群は真面目グループと似た回答であるが、終盤群になるほど回答時に慣れが生じ、不真面目グループと似た回答となっていると考えられる。そのため、リズム化グループの実験データが、真面目グループの実験データと区別せずに扱われると、実験データの分析の質を下げってしまう恐れがある。

同じ選択位置を 3 回以上連続して回答した割合について、データセット A, B, C いずれの結果においても、リズム化グループは序盤では同じ選択肢を連続して選択しておらず、終盤に同じ選択肢を連続して選択する傾向が観測された。このことから、序盤は真面目に選択しているものの、終盤には不真面目となり、不適切な回答をしているものと考えられる。なお、データセット A, B においては、不真面目

グループは3回以上連続して同じ選択肢の位置を選ぶ傾向が高く、不適切な回答をしていることが見て取れる。なお、データセットCにおいてその傾向が現れなかったのは、2択と3択が切り替わるためであると考えられる。

以上の結果より、「回答作業中に慣れが生じてしまい、回答がリズム化した実験協力者は回答時間も短くなり、回答も不適切になる」という仮説が立証された。

今回の実験では、単純な選択に関する実験を扱ったが、単なる選択だけでなくチェックボックスやラジオボタンをクリックして選択したり、マウスで対象をひたすらクリックしたり、自由記述に回答したりといったような様々な操作が要求されるようなタスクにおいても、リズム化が生じる可能性がある。そこで今後は、こういった操作であればリズム化が生じやすいのかについて検証を行っていく予定である。また、今回は繰り返しの実験を対象としたが、大規模なデータセット構築などにおいて、与えられた設問に対して回答の理由を自由記述でひたすら回答していくものや、写真やフォントなどを見て該当する印象にまつわる形容詞対にひたすら回答を行うようなものにおいても、リズム化による慣れが悪い影響を及ぼす可能性がある。そこで今後は、これらのより複雑な条件についてもリズム化による慣れの影響が生じているか、またこういった問題があるかについて調査を行う予定である。

7. まとめと展望

本研究では、Web上でのアンケート調査や実験、データセット構築などにおいて生じる慣れの問題について、我々の研究室でこれまで実施してきた繰り返し選択実験[4][11][12]をもとに、「回答作業中に慣れが生じてしまい、回答がリズム化した実験協力者は回答時間も短くなり、回答も不適切になる」という仮説を立て、実験データについて多面的に分析を行った。ここでは、実験者を真面目、リズム化、不真面目の3グループに分類し、グループごとに設問群および設問ごとに回答時間の分析を行った。その結果、分析を行ったすべてのデータセットで、真面目、リズム化、不真面目の順に回答時間が長いことが明らかになった。また、リズム化グループは設問の前半部分では回答時間が長く、真面目グループに似ている傾向があるものの、後半に近づくると回答時間が減少し不真面目グループの結果に近くなるという傾向が見られた。このことから、回答時間に着目した分析は、実験者の状態をある程度判定することが可能であり、慣れの影響を検出する要因として有効であると考えられる。さらに、各グループにおける連続回答の回数についても分析を行った。その結果、全てのデータセットでリズム化グループの同じ選択肢の連続回答の回数が設問群ごとに増加しており、慣れによる問題が確認された。

今後の研究として選択肢の連続回答が慣れによる影響で生じるかについて、実験環境を整えて再度データセットを構築し、より詳細な分析を行うとともに、回答時間で慣れの影響を検出するシステムを実現する予定である。また、設問に対して自由記述で回答を行う実験や、写真やフォントを見て該当する形容詞対の回答を行う印象評価データセット構築などのより複雑な作業でも、慣れの影響が生じているかについて調査を行う予定である。さらに、今回はリズム化による慣れに関する基礎的な調査を行ってきたが、今後はこうしたリズム化による慣れを検知し、防止する手法も検討予定である。

謝辞 本研究の一部は科研費 JP22K12135 の助成を受けたものです。また実験データの構造および分析についてアドバイスいただいた木下裕一朗さん、関口祐豊さん、植木里帆さんに深く感謝します。

参考文献

- 1) “Yahoo!クラウドソーシング”. <https://crowdsourcing.yahoo.co.jp/>, (参照: 2022-12-14).
- 2) “Amazon Mechanical Turk”. <https://www.mturk.com/>, (参照: 2022-12-14).
- 3) 横山幸大, 中村聡史. 待機画面におけるプログレスバーの表示位置が待機後の選択に及ぼす影響の調査. 情報処理学会研究報告ヒューマンコンピュータインタラクション (HCI), 2022, vol. 2022-HCI-200, no. 1, pp. 1-8, 2022.
- 4) 木下裕一朗, 関口祐豊, 植木里帆, 横山幸大, 中村聡史. 選択インタフェースにおけるアイテムの遅延表示が選択に及ぼす影響. 情報処理学会 研究報告ヒューマンコンピュータインタラクション (HCI), 2022, vol. 2022-HCI-200, no. 27, pp. 1-8.
- 5) 三浦麻子, 小林哲郎. オンライン調査モニタの Satisfice 行動に関する実験的研究. 社会心理学研究, 2015, vol. 31, no. 1, p. 1-12.
- 6) 山崎郁未, 伊藤理紗, 中村聡史, 小松孝徳. Web アンケートにおける不真面目回答予防システム実現に向けた自由記述配置の基礎検討. 情報処理学会 研究報告ヒューマンコンピュータインタラクション (HCI), 2021, vol. 2021-HCI-195, no. 34, pp. 1-8.
- 7) Rankin, C. H., Abrams, T., Barry, R. J., Bhatnagar, S., Clayton, D. F., Colombo, J., Coppola, G., Geyer, M. A., Glanzman, D. L., Marsland, S., McSweeney, F. K., Wilson, D. A., Wu, C-F., Thompson, R. F. Habituation Revisited: An Updated and Revised Description of the Behavioral Characteristics of Habituation. *Neurobiology of Learning and Memory*, 2009, vol. 92, no. 2, pp. 135-138.
- 8) Howarth, P. A., Hodder, S. G.. Characteristics of habituation to motion in a virtual environment. *Displays*, 2008, vol. 29, no. 2, pp. 117-123.
- 9) 坂田大, 早川太揮. 快情動喚起と脱馴化でドライバの覚醒状態を維持する手法の開発. 自動車技術会論文集, 2021, vol. 52, no. 3, pp. 589-594.
- 10) 中野泰彦, 小島孝雄, 河中治樹, 小栗宏次. 指差し呼称を用いた高齢者ドライバー交通事故削減のための注意力向上方法の検討. 電子情報通信学会 論文誌 D, 2014, vol. J97-D, no. 1, pp. 135-144
- 11) 植木里帆, 横山幸大, 野中滉介, 中村聡史. 三択の選択肢における要因の違いが選択行動に及ぼす影響の調査. 情報処理学会 研究報告ヒューマンコンピュータインタラクション

- (HCI) , 2020, vol. 2020-HCI-190, no. 23, pp. 1-8.
- 12) 関口祐豊, 植木里帆, 横山幸大, 中村聡史. 三択の選択肢の色の組み合わせが選択行動に及ぼす影響. 情報処理学会 研究報告ヒューマンコンピュータインタラクション (HCI) , 2021, vol. 2021-HCI-195, no. 32, pp. 1-8.
 - 13) Cai, L., Zhu, Y.. The Challenges of Data Quality and Data Quality Assessment in the Big Data Era. *Data Science Journal*, 2015, vol. 14, no. 2, pp. 1-10.
 - 14) Chawla, N. V., Japkowicz, N., Kotcz, A.. Editorial: Special Issue on Learning from Imbalanced Data Sets. *Association for Computing Machinery*, 2004, vol. 6, no. 1, pp. 1-6.
 - 15) Lateh, M. A., Muda, A. K., Mohd, Z, I, Y., Muda, N. A., Azmi, M. S.. Handling a Small Dataset Problem in Prediction Model by employ Artificial Data Generation Approach: A Review. *Journal of Physics*, 2017, vol. 892, no. 012016, pp. 1-10.
 - 16) 松井直, 郭清蓮, 北村大輔. 教師データ収集支援システムに関する研究と開発. 情報処理学会, 研究報告コンピュータビジョンとイメージメディア (CVIM) , 2019, vol. 2019-CVIM-215, no. 30, pp. 1-4.
 - 17) 溝口耕三, 岡本健久, 田中洪. 選択的注意から見た「慣れ」の測定法の研究. *日本音響学会誌*, 1999, vol. 55, no. 5, pp. 343-350
 - 18) 菊池英明, 北村泰司, 本島嘉朗, 山田哲史, 宮沢幸希, 白勢彩子. 音声インタフェース評価における慣れの影響の分析. 情報処理学会, 情報処理学会研究報告音声言語情報処理 (SLP) , 2007, vol. 2007-SLP-67, no. 75, pp. 97-102.
 - 19) 高田友樹, 笹倉万里子. インタラクションにおける慣れの効果の調査. 情報処理学会 インタラクション 2020 論文集, 2020, 3B-57, pp. 992-997.
 - 20) 杉邑洋樹, 上野秀剛. 視線を利用したユーザインタフェースに対するユーザの慣れの定量化. 情報処理学会, 研究報告ヒューマンコンピュータインタラクション (HCI) , 2012, vol. 2012-HCI-150, no. 9, pp. 1-7.
 - 21) Wood, D. D., Fisher, D. L., Andres, R. O.. Minimizing Fatigue during Repetitive Jobs: Optimal Work-Rest Schedules. *Human Factors*, 1997, vol. 39, no. 1, pp. 83-101.
 - 22) Smith, P. C., Lem, C.. Positive aspects pf motivation in repetitive work: effects of lot size upon spacing of voluntary work stoppages. *Journal of Applied Psychology*, 1955, vol. 39, no. 5, pp. 330-333.
 - 23) 辛島光彦, 西口宏美. 単純繰り返し作業における作業前音楽聴取の有効性に関する研究: 転記作業と心的回転作業を例に. 2012, vol. 63, no. 2, pp. 29-40.