

推薦システムにおけるリスト型とマップ型の 提示インタフェースと推薦誘導への気づきに関する心理学実験

竹内 鴻[†] 福島 優都[†] 佃 洸^{††} 濱崎 雅弘^{††} 後藤 真孝^{††} 土方 嘉徳^{†,††}

[†]関西学院大学大学 商学部 〒662-8501 兵庫県西宮市上ヶ原一番町 1-155

^{††}産業技術総合研究所 〒305-8568 茨城県つくば市梅園 1-1-1 中央第二

^{†††}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 丁目 2-1

E-mail: ^{†††}contact@soc-research.org

あらまし 本研究では、推薦システムのインタフェース (IF) の違いが推薦誘導 (特定の商品やコンテンツを選択するようにユーザを誘導すること) への気づきやすさとユーザビリティに及ぼす影響を検証した。楽曲推薦機能を有する音楽発掘サービス「Kiite」(リスト型 IF) とその一機能である「Kiite World」(マップ型 IF) を用い、35 名を対象に心理学実験を行った。その結果、推薦誘導の気づきに IF 間で統計的有意差は確認されなかったが、観察されたデータではマップ型で正しく気づいた参加者がわずかに多かった。一方、ユーザビリティの評価は全体としてリスト型が高かったが、一部の参加者はマップ型を高く評価することも確認された。

キーワード 推薦システム, インタフェース, リスト型, マップ型, 推薦誘導, ユーザビリティ

1. はじめに

近年、推薦システムは楽曲、動画の配信サイトやニュースサイトなどに幅広く導入されている。しかし推薦システムを含んだ意思決定支援システムを信頼できるかどうかを結果や過程の説明などから判断することは難しく [1], 結果を無批判に受け入れれば、それが意図的に操作された場合に、ユーザの消費行動や視聴行動そのものが誘導されてしまう危険性がある [2]。本研究ではこのような操作による誘導を「推薦誘導」と呼ぶ。研究者によっては、そうした操作を「ポイズニング攻撃(poisoning attack)」と呼ぶこともある。Wang らは、ポイズニング攻撃を「推薦システムの学習データに悪意あるデータを注入し、モデルの学習過程を操作して推薦結果を攻撃者の目的に沿って歪める行為」と定義している [3]。このような推薦の操作を受けると、ユーザはそれに気づかずに推薦された商品やコンテンツを消費・閲覧してしまうかもしれない。しかし、推薦誘導に対してユーザが気づけるかどうかは、これまで十分に検討されてこなかった。

本研究では、推薦結果の提示方法、すなわちインタフェース (IF) の形式によっては、ユーザの推薦誘導への気づきやすさに違いが生じると考える。既存サービスで一般的な「リスト型の IF」は、ユーザへのお薦め度合いの高い順に縦方向にアイテム (楽曲や映画などの推薦システムで提示される対象) が表示されることが多く、上位に注意が集中しやすいことが分かっている [4]。しかし順序以外の情報に乏しく、推薦誘導が行われたとしても、ユーザはそれに気づきにくいと予想される。

これまで推薦の IF の形式には、リスト型・マップ型・円型などが提案されている [4, 5]。なかでも「マップ型の IF」では、アイテムを 2 次元または 3 次元空間に配置して、推薦結果を提示する IF である。これによりユーザは、アイテム間の位置関係や距離を考慮することができるようになるため、アイテム間の類似性や関係性を直感的に把握できるようになる。そのため、ユーザは推薦誘導の存在に気づきやすくなると考えら

れる。本研究では順序情報に依存するリスト型と、順序以外の情報 (距離や空間位置) が反映されるマップ型の 2 つの IF を取り上げ、推薦誘導の気づきやすさに違いがあるかどうかを明らかにする。

IF の違いが推薦誘導の気づきやすさに影響を与えるかを確かめるためには、実際にユーザに推薦システムを使用してもらい、そこで推薦誘導を行って確かめる必要がある。すなわちユーザに対する心理学実験で明らかにする必要がある。この目的を達成するためには、リスト型とマップ型の両方の IF を備えた推薦システムを利用することが理想的である。そのような Web サービスとして、第 3~5 著者が研究開発をしながら一般公開している、楽曲推薦機能を有する音楽発掘サービス「Kiite¹」とその一機能である「Kiite World²」がある。Kiite ではリスト型の IF でユーザの興味や嗜好にあった歌声合成楽曲が提示される [6]。Kiite World では Kiite と同じ推薦結果が 2 次元空間上に提示される (以降、その推薦結果が提示される 2 次元平面を「マップ」と呼ぶ)。また Kiite World では、各ユーザの座標の位置もユーザの「家」としてマップ上に可視化され、ユーザの家の近くには、そのユーザが好きそうな楽曲 (つまり、推薦スコアの高い楽曲) が配置される [7]。さらに、各楽曲の座標位置は全ユーザのマップで共通であるため、推薦する楽曲の提示位置を一人ひとりに合わせて改変することはできない。そのため、仮にあるユーザに興味のない楽曲を推薦すると、その楽曲はユーザの家から離れた位置に表示されるため、ユーザは推薦誘導に気づく可能性が高いと考えられる。

ただし、推薦誘導に気づきやすかったとしても、推薦システム全体の体験が損なわれると、ユーザはその推薦システムを継続して利用しなくなる可能性がある。そのため本稿では、IF が使いやすいかどうか、すなわちユーザビリティについても検証する。これまでリスト型やグリッド型などの様々な IF におけるユーザビリティの比較研究は存在するものの [4, 8], 著者らの調査した限りでは、リスト型とマップ型のユーザビリティを直接比較した研究は存在しない。また、IF の使いやすさだけではなく、態度 (心理学の分野で、特定

¹ <https://kiite.jp/>

² <https://world.kiite.jp/>

の対象や事柄に対して、人が比較的持続的に抱く好意的または否定的な評価（心理的傾向）のこ）を向上させているのかも検証されていない。

以上を踏まえ、本研究では以下の Research Question (RQ) を設定する。

RQ1：推薦誘導が行われたとき、リスト型とマップ型のどちらの IF の方がユーザは気づきやすいか？

RQ2：リスト型とマップ型のどちらの IF の方が、ユーザビリティとユーザの態度は高いか？

本研究では上記の RQ を心理学実験で明らかにするために、実験のドメインとして音楽を採用し、推薦システムにはリスト型 IF として「Kiite」を、マップ型 IF として「Kiite World」を採用する。

2. 方法

本研究では、IF の違いによる推薦誘導への気づきやすさとユーザビリティの違いを検証するための研究方法として、実際に実験参加者に推薦システムを使って評価してもらう心理学実験を採用した。また、実験結果の分析には、定量分析と定性分析を組み合わせる混合研究法 (mixed methods research) を採用した[9]。定量分析では、推薦誘導に気づいた参加者の割合やユーザビリティ指標の回答値を評価する。一方で、「どのような要因によって推薦誘導に気づいたのか」は、事前に網羅的な質問項目を設定して定量的に測定することが難しい。そのため、本研究では自由記述により推薦誘導に気づいた理由を尋ね、定性分析を通じて、気づきの要因を探索する。

2.1. 実験対象

調査会社の協力で参加者を募集した。参加条件は、(1) 音楽発掘サービス「Kiite」の利用経験がないこと、(2) 歌声合成楽曲への肯定的な関心を持っていることの 2 点を設定した。最終的に 35 名の参加者が調査に協力した。実験を最後まで完了した参加者には謝礼として現金 2 万円を支払った。また、実験中に提示された楽曲のうち、参加者が気に入ったものについては、楽曲名とクリエイタ名の情報をファイルに出力して持ち帰り、実験後にも聴くことができるようにした。これは、参加意欲を高めるための工夫の一つである。

2.2. 実験方法

本研究では、音楽発掘サービス「Kiite」および「Kiite World」を実験に用いた。Kiite では各ユーザの嗜好に基づいて楽曲を推薦する機能を提供しており、推薦結果はリスト型で提示される[6]。一方 Kiite World は、Kiite の推薦の仕組みを用いて推薦結果をマップ型で提示する[7]。具体的には、Kiite の推薦モデルで高次元ベクトルで表現されるすべてのユーザとすべての楽曲を、UMAP [10]に基づいて 2 次元ベクトルに非線形に変換し、2 次元ベクトルを座標値とみなして 2 次元平面上にユーザと楽曲を配置する。このとき、各ユーザの座標値は「家」としてマップ上に可視化され、その周辺には当該ユーザの嗜好に合致した楽曲群が配置される。これにより、ユーザは自分の家を起点にマップを探索し、好みの楽曲を効率的に発見することが可能となる。ただし、これらの Web サービスは一般公開中でユーザに活用されており、画面上には本実験とは関係しない多くの機能が表示されている。そこで、第 3～5 著者が「Kiite」および「Kiite World」を研究開発している強みを活かし、本実験のために表示内容をシンプルにした専用モードを新規に開発して用いた（以下で例示する図 1, 2, 3 は、すべて専用モードの画面例と

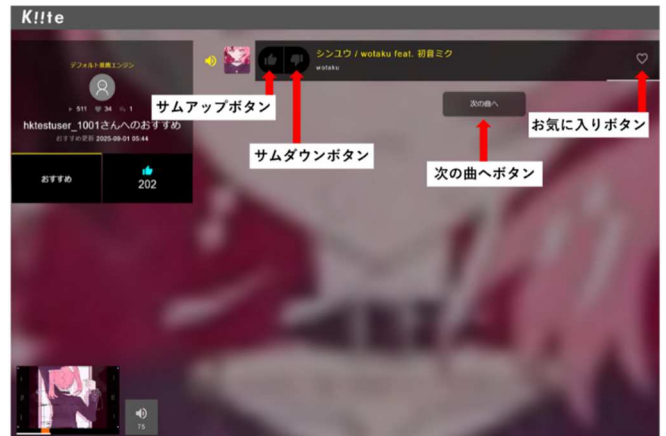


図 1 Kiite の学習フェーズの画面例

なり、推薦に関する機能は同一でも画面は一般公開されているサービスとは異なる）。

また、これらのサービスの推薦モデルは共通だが、ユーザに適切な楽曲を推薦するためには事前に個人の嗜好に関するデータが必要である。ユーザが Kiite および Kiite World を初めて使用する場合、Kiite World のマップ上にそのユーザの家が表示されるのは、嗜好に関するデータが得られた翌日になるため、実験は以下に述べるように 2 日間に分けて実施する形式を採用した。

1 日目（「学習フェーズ」と呼ぶ）は、参加者の嗜好に関するデータを収集するため、Kiite に嗜好データ収集用の専用モードを本実験用に開発し、その上で参加者は提示された各楽曲を聞いて判断した嗜好を「サムアップ」または「サムダウン」のボタンを押して評価した（図 1 参照）。評価は必須ではなく、サムアップとサムダウンのいずれのボタンも押さずに「次の曲へ」ボタンを押すと、直ちに次の楽曲が提示され再生される。また、気に入った楽曲があれば、「お気に入り」ボタンを押すことで、2.1 節で述べたように実験後にその楽曲のデータを持ち帰ることができる。このフェーズでは、実験実施者による説明の時間が 10 分で、その後の嗜好データ収集の時間が 60 分であった。

2 日目は第 1～4 ステップで構成され、第 1・2 ステップを通常推薦フェーズ、第 3・4 ステップを推薦誘導フェーズと呼ぶ。参加者は、各ステップでリスト型とマップ型のインタフェースを交互に利用した（図 2, 3 参照）。各ステップは、当該ステップの実験説明（10 分）と、実験参加（インタフェースを利用した楽曲の探索）（20 分）、質問紙調査（10～15 分）で構成される。各ステップの間にはおよそ 20 分の休憩を設けた。各ステップの説明では、システムの操作方法に加え、楽曲を聴く順番や 1 曲あたりの聴取時間に制限がないことを参加者に伝えた。

Kiite World では、最初に参加者のアイコンが参加者の家の位置に表示される。参加者はマップ上で自分のアイコンを自由に動かしながら、楽曲のサムネイルに触れることで楽曲を聴くことができる。また、図 3 の左下の「ミニマップ」には、推薦されたすべての楽曲のマップ上の位置が表示される。ミニマップをクリックすると拡大して表示され、その後ミニマップ上の任意の位置をクリックすると、その位置に瞬時に自分のアイコンを移動させることができる。

第 1 ステップで提示される推薦結果は学習フェーズで収集した嗜好データをもとに生成され、第 2 ステッ



図 2 Kiite (リスト型インタフェース) の画面例

プ以降は、直前のステップにおける推薦結果への評価に基づいて更新された推薦結果が提示された。

順序効果を考慮し、実験はリスト型から開始する群 (A 群) とマップ型から開始する群 (B 群) に分けた。さらに第 3・4 ステップは参加者ごとに「誘導あり」または「誘導なし」の条件に割り当てられ、両ステップとも誘導の有無以外は同じ条件で実施した。また、第 3・4 ステップでは、すべての参加者に対し、「推薦結果に本来推薦すべきでない楽曲が混入している可能性がある」旨を事前に説明した上でタスクに取り組ませた。結果として、「A 群誘導あり」「A 群誘導なし」「B 群誘導あり」「B 群誘導なし」の 4 群を設定した。後の分析では、誘導ありの場合と誘導なしの場合に分けて分析を行うため、A 群誘導ありと B 群誘導ありをまとめて「誘導あり群」、A 群誘導なしと B 群誘導なしをまとめて「誘導なし群」と呼ぶ。なお、2 日目の実験環境は統制し、ヘッドホン (オーディオテクニカ社製 ATH-MSR7b, ATH-M20x) とノート PC (Lenovo 社製 14 インチ, 1920×1080) を使用した。

参加者 *u* に対して推薦誘導を行う楽曲 (誘導曲) は、以下の手順で選択した。

- (1) *u* が実験に参加した日から遡って過去 1 年間に投稿された楽曲を選択する。
- (2) (1) の楽曲の中から、Kiite 上のお気に入り数が 5 件以上 15 件以下の楽曲を選択する。
- (3) (2) の楽曲の中から、1 日目の *u* への推薦結果および、2 日目に実施済みのステップでの *u* への推薦結果に含まれていない楽曲を選択する。
- (4) (3) の楽曲の中から、Kiite World のマップ上で *u* の家の座標から遠い順に上位 20 曲のランキングを作成する。
- (5) (4) のランキングの奇数 (または偶数) 順位の楽曲をリスト型用の誘導曲、偶数 (または奇数) 順位の楽曲をマップ型用の誘導曲として選択する。

(1) および (2) により、比較的新しく、かつクオリティがある程度担保されている楽曲を選択している。さらに (3) により、参加者が実験中に目にしていない楽曲を選択している。(4) では、参加者の嗜好に合致しない楽曲を誘導曲として選択している。(5) で奇数順位と偶数順位のどちらをリスト型の誘導に用いるかは、参加者間で均等になるように決めた。もし仮にプラットフォームが推薦誘導を行う際は、ユーザの嗜好とは無関係に、比較的新しいアイテムの中から利益率の高いアイテムや人気を獲得させたいアイテムへ誘導することが



図 3 Kiite World (マップ型インタフェース) の画面例

考えられる。本実験では、そうした推薦誘導を擬似的に実現するために、上記のような手順により、推薦誘導を行うための楽曲を選択した。

2.3. 評価尺度

推薦誘導に気づいたかどうかに関しては、推薦誘導が行われていたと思ったかどうかを第 4 ステップの後に質問した。質問に対して「行われていない」「どちらかといえば行われていない」「どちらかといえば行われている」「行われている」の 4 段階で回答してもらい、自由記述でその理由も回答させた。ユーザビリティの評価には、O'Brien らのユーザーエンゲージメント尺度 (User Engagement Scale)[11] を本実験に合うように修正した 4 下位尺度各 3 項目を用いた。これらは 5 段階のリッカート尺度で測定した。下位尺度は、FA

(Focused Attention: 対話に没頭し時間の経過を忘れる感覚を表す)、PU (Perceived Usability: 対話の結果として経験したネガティブな感情、および制御感と費やした努力の度合いを表す)、AE (Aesthetic Appeal: IF の魅力と視覚的訴求力を表す)、RW (Reward: 価値ある経験的成果を表す) で構成される。

さらに、ユーザビリティに関連して、システムの総合的な利便性を尋ねる質問として、TAM (Technology Acceptance Model)[12]における TAM-PU (知覚的有用性) (7 項目) と TAM-PEOU (知覚的使いやすさ) (4 項目) を導入し、前者は Armentano の質問文[13]、後者は Armentano の質問文[13]、Pu と Chen の質問文[14]、Lengyel の質問文[15]を本実験に合うように修正して利用した。これらは 7 段階のリッカート尺度で測定した。

また、IF へのユーザの態度 (心理学における中心的な心的概念であり、ある対象に対して比較的持続的に抱く、評価的な反応傾向を表す) として、推薦システムの信頼性研究でもよく用いられている継続利用意図 (続けて対象の製品やサービスを使い続ける意図) (5 項目)、満足度 (対象の製品やサービスへの総合的な評価) (4 項目)、信頼 (対象の製品やサービスへの信頼) (9 項目) の概念を尋ねた。継続利用意図は Hsu と Lu

の尺度[16]、Benbasat と Wang の尺度[17]、態度は Flavián らの尺度[18]、信頼は Lee と Choi の尺度[19]で提案された心理尺度を本実験に合うように修正して用いた。これらは 7 段階のリッカート尺度で測定した。

上記のユーザビリティに関する調査は、第 1~4 ステップの各ステップで実施した。ユーザビリティについて分析する際は、各質問の点数の合計を質問数で割ることで標準化を行った。

上記に加えて、年齢、性別、最終学歴をデモグラフィック情報として取得した。また、推薦誘導を行った群に対しては、自分の家から離れた場所に提示された楽曲の評価や聴取行動に関する質問8問を設定し、そのうち一部には詳細を自由記述で回答させた。さらに、楽曲選択時に重視した提示情報（リスト内の楽曲の順位、マップ上の楽曲の位置、サムネイル、タイトル、クリエイタ名）や聴取をためらった楽曲の理由についても計8問で尋ねた。後者の8問の項目については、本論文では分析対象とはせず、今後の研究で取り組む。

2.4. 分析方法

定量分析では、リスト型とマップ型のIF間の誘導への気づきやすさに違いがあるかどうかの検証と、IF間でユーザビリティに違いがあるかどうかの検証をするために、t検定を主とする2群の平均値の差の検定を行った。分析ソフトウェアはRを用いた。1人の参加者において、学習フェーズで不手際があったため、その者は分析対象から除外した。t検定に先立ち正規性を確認した。定性分析については、自由記述の内容で代表的なものを研究者らで確認を行った。

3. 結果

3.1. 推薦誘導への気づき

3.1.1 定量分析

推薦誘導への気づきについて4段階で収集した回答を2段階に集約した。具体的には、「どちらかといえば行われていたと思う」「行われていたと思う」「行われていると思った」群に、「行われていなかったと思う」「どちらかといえば行われていなかったと思う」「行われていないと思った」群に統合した。行われていると思った群／行われていないと思った群の人数をリスト型とマップ型で集計し、誘導なし群と誘導あり群のそれぞれについてカイ二乗検定を行った。その結果、誘導あり群では $\chi^2(1) = 0.118$, $p = 0.731$ 、誘導なし群では $\chi^2(1) = 0.113$, $p = 0.737$ となり、いずれにおいてもIF間で有意差は確認されなかった。

しかし傾向としては、推薦誘導あり群の場合には、図4および表1に示すように、マップ型の方が推薦誘導が「4: 行われている」と回答した参加者がリスト型よりも多く、逆に「1: 行われていない」と回答した参加者は少なかった。一方、推薦誘導なし群の場合には、図5および表1に示すように、マップ型の方が推薦誘導が「1: 行われていない」と回答した参加者が多く、「4: 行われている」と回答した参加者は少なかった。このようにマップ型では、推薦誘導の有無のいずれの場合でも、実際の挙動と一致する方向に気づけている回答がより多い傾向がみられた。

以上から、参加者数の少なさから統計的有意差を得ることはできなかったが、推薦誘導の有無に限らず、マップ型の方が推薦誘導が行われているかどうかを正しく判断できる可能性が示された。

3.1.2 定性分析

推薦誘導の有無に気づいた理由についての自由記述を分析したところ、マップ型では数は限られるが、アイテムの表示位置に着目した回答が散見された。例えば以下のような記述が得られた。

「1回目のマップには表示されていない遠いエリアに数曲配置されていたこと、それが特に興味があるわけでもなかったから。MVに流れてくるコメント数が極端に少ないものがあり、マイナーな

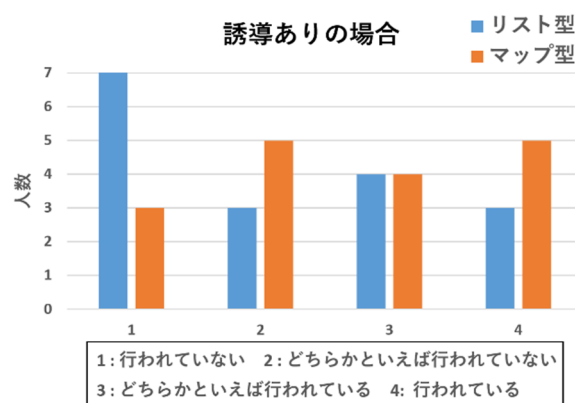


図4 推薦誘導に気づいたかを問う質問への回答の分布（推薦誘導あり群）

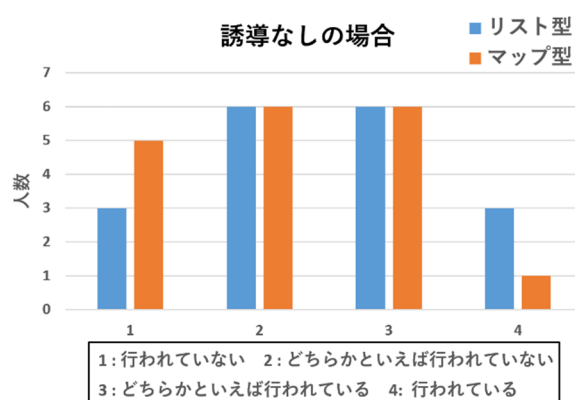


図5 推薦誘導に気づいたかを問う質問への回答の分布（推薦誘導なし群）

表1 クロス集計表

	推薦誘導あり群	
	行われていると思った群	行われていないと思った群
リスト型	7	10
マップ型	9	8

	推薦誘導なし群	
	行われていると思った群	行われていないと思った群
リスト型	9	9
マップ型	7	11

曲をプッシュしたいのかなと感じたから」(誘導あり群、推薦誘導に気づいた：参加者番号31)

「特に嫌いな曲が表示されなかった。マップの家から遠い曲がいくつか表示されているのは気になった」(誘導あり群、推薦誘導に気づかなかった：参加者番号24)

一方で、リスト型ではリスト内の楽曲の位置(順位)に基づいた言及はほとんど見られず、多くの参加者は実際に楽曲を聴き、自身の嗜好やジャンルとの一致・不一致を根拠として判断していた。例えば以下のよう

表 2 群間の比較の結果

尺度	リスト型(M, SD)	マップ型(M, SD)	統計量	p 値	Hedges' g
PU	4.157, 1.008	3.222, 0.997	$t(32.840)=2.757$	9.45e-3*	0.911
TAM-PU	5.555, 0.764	4.508, 1.417	$t(26.420)=2.740$	0.0108*	0.891
TAM-PEOU	6.044, 0.741	4.861, 1.443	$t(25.685)=3.075$	4.93e-3*	0.999

* $p < .05$

な記述が得られた。

「違和感を感じるような、特別好みとかけ離れたような曲は流れてこなかったように思えるから」
(誘導あり群、推薦誘導に気づかなかった：参加者番号 31)

3.2. マップ型・リスト型のユーザビリティの違い

推薦誘導が行われていない通常時の推薦システムの利用の状況の方が、ユーザが経験する機会が多いと考えたため、推薦誘導が行われていない推薦フェーズ（つまり、第 1 ステップおよび第 2 ステップ）を対象として、マップ型とリスト型のどちらがユーザビリティが高いかを検証した。

それぞれの尺度は、リッカート尺度で複数の質問によって収集したため、分析方法は、Norman の主張を参考にパラメトリック検定で行うことにする[20]。そこで、群間の比較には、等分散性の前提を必要としない Welch の t 検定を採用した。

第 1 ステップと第 2 ステップの各ステップにおいてユーザビリティに関する質問紙調査を行っているが、参加者はそれぞれのステップで異なる IF を用いている。つまり第 1 ステップでは、A 群と B 群の参加者はそれぞれ初めて本実験用のリスト型 IF とマップ型 IF を利用することになり、参加者は比較対象がない中で、対象の IF の評価を行う必要があった。そのため、参加者が「様子見」として比較的的中した評価値を付ける傾向があると我々は考えた。第 1 ステップと第 2 ステップの両方のデータを用いれば、実験参加者内計画として分析することはできるが、どちらの IF から使用するかにより、同じステップにおいても評価値は異なってくると思われる。そこで、相対比較が可能になった第 2 ステップのデータの方が信頼性が高いと考え、第 2 ステップのデータを用いて、実験参加者間比較としてユーザビリティの評価を行った。すなわち、A 群の第 2 ステップ（マップ型）の評価値と B 群の第 2 ステップ（リスト型）の評価値を比較することにした。

上記の分析の結果、表 2 に示すように、PU、TAM-PU、TAM-PEOU において有意にリスト型の評価が高かった。全体としてはマップ型よりもリスト型の評価が高い傾向が見られたが、必ずしもすべての参加者がリスト型を高く評価していたわけではなかった。

そこで参加者ごとの評価値の差を確認するため、ユーザビリティの各尺度について差分スコア（マップ型の評価値からリスト型の評価値を減じたもの）を算出した。尺度ごとに 5 段階または 7 段階で得た評価の合計をポイント化し（この評価では標準化は行わなかった）、項目数が 3, 4, 5 個のものは 2 ポイント、7, 9 個のものは 3 ポイント以上高くマップ型を評価していた場合、その尺度ではマップ型を高く評価したとみなした。そのうえで、マップ型を高く評価した尺度の数を各参加者について集計した。リスト型についても同様に集計した。差分が閾値未満の項目については、評価

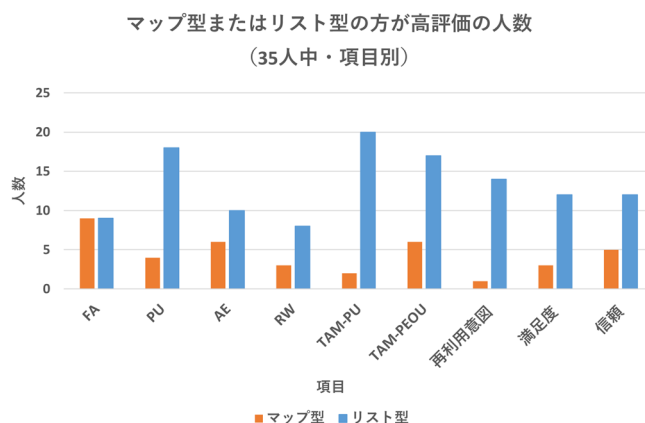


図 6 マップ型またはリスト型の方が高評価の人数

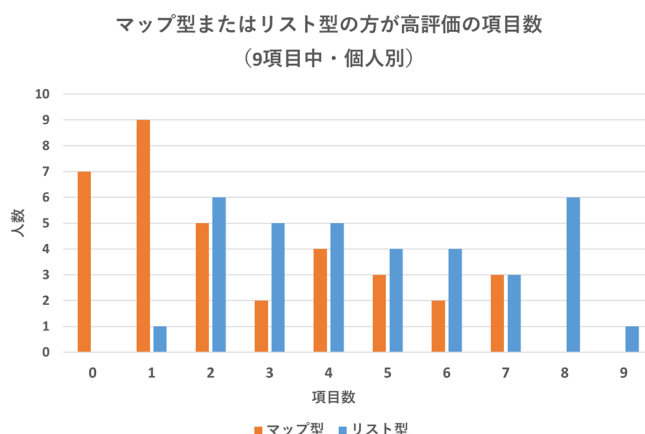


図 7 マップ型またはリスト型の方が高評価の項目数

に大きな違いはないものとして除外した。

集計結果を図 6 に示す。集計の結果、FA・TAM-PEOU・満足度など全 9 尺度のいずれにおいても、程度の差はあるもののマップ型を高く評価する参加者が一定数存在することが確認された。また、図 7 に示すように、9 尺度のうち過半数（5 項目以上）でマップ型を高く評価した参加者は 35 人中 8 人であった。これらのことから、全体としては既存サービスで目にする機会が多いリスト型が高く評価される一方で、マップ型を優位に評価する参加者も一定数存在し、IF 評価には個人差がある可能性が示唆された。

3.3. 研究の制約

本研究ではサンプル数が限られていたため、推薦誘導への気づきとユーザビリティのいずれの定量分析においても、有意差が確認された尺度は一部にとどまっ

た．今後データ数を増やすことで，リスト型とマップ型の推薦誘導への気づきやすさやユーザビリティの違いをより明確にできると期待できる．ユーザビリティの評価は，全体平均ではリスト型の方が高かったが，評価は個人差も大きく，参加者によってはマップ型を高く評価する場合もあった．こうした評価のばらつきがあるため，ユーザビリティの評価を単純な平均値の比較だけではなく，どのような目的で当該サービスを使っているかなど，より詳細なユーザの性質や状況も考慮して分析する必要があると考える．

4. おわりに

本研究では，マップ型とリスト型の IF を比較し，推薦誘導への気づきやすさやユーザビリティの評価との関係を検討した．その結果，推薦誘導への気づきに関して統計的有意差は確認されなかったが，観察されたデータではマップ型 IF を用いることで正確に判断できた参加者がわずかに多かった．ユーザビリティの評価では平均的にはリスト型が高かった一方，参加者によってはマップ型をより高く評価する例もみられた．

今後は，ユーザビリティにおいてマップ型を高く評価するユーザが推薦誘導に気づきやすい傾向を持つかを検証するとともに，推薦誘導への気づきやすさと推薦システムへの態度との関係を分析する予定である．

謝辞

本研究の一部は JST CREST JPMJCR20D4 の支援を受けた．

文 献

- [1] I. Nunes and D. Jannach, "A systematic review and taxonomy of explanations in decision support and recommender systems," *User Modeling and User-Adapted Interaction*, vol.27, no.3-5, pp.393-444, Dordrecht, The Netherlands, May-July 2017.
- [2] T. T. Nguyen, Q. V. H. Nguyen, T. T. Nguyen, T. T. Huynh, T. T. Nguyen, M. Weidlich, and H. Yin, "Manipulating recommender systems: A survey of poisoning attacks and countermeasures," *arXiv preprint, arXiv:2404.14942*, no.1, pp.1-24, Apr.2024.
- [3] Z. Wang, M. Gao, J. Yu, H. Ma, H. Yin, and S. Sadiq, "Poisoning attacks against recommender systems: A survey," *arXiv preprint, arXiv:2401.01527*, no.1, pp.1-9, Jan.2024.
- [4] L. Chen and H. K. Tsoi, "Users' decision behavior in recommender interfaces: Impact of layout design," *Joint Proc. RecSys 2011 Workshop on Human Decision Making in Recommender Systems (Decisions@RecSys'11) and User-Centric Evaluation of Recommender Systems and Their Interfaces-2 (UCERSTI 2)*, no.4, pp.21-26, Chicago, IL, USA, Oct.2011.
- [5] P. Knees, M. Schedl, and M. Goto, "Intelligent user interfaces for music discovery," *Transactions of the International Society for Music Information Retrieval*, vol.3, no.1, pp.165-179, Oct.2020.
- [6] 佃洗撰, 石田啓介, 高橋卓見, 濱崎雅弘, 後藤真孝, "レイヤー構造に基づく楽曲推薦手法の提案と音楽発掘サービス Kiite への応用," *情報処理学会研究報告 音楽情報科学 (MUS)*, vol.2025-MUS-142, no.11, pp.1-12, Feb.2025.
- [7] K. Tsukuda, T. Takahashi, K. Ishida, M. Hamasaki, and M. Goto, "Kiite World: Socializing Map-Based Music Exploration Through Playlist Sharing and Synchronized Listening," *Proc. 31st International Conference on MultiMedia Modeling (MMM 2025)*, *Lecture Notes in Computer Science*, vol.15521, pp.197-211, Springer, Singapore, Jan. 2025.
- [8] D. Jannach, M. Jugovac, M. Jesse, and C. Trattner, "Exploring Multi-List User Interfaces for Similar-Item Recommendations," *Proc. 29th ACM Conf. on User Modeling, Adaptation and Personalization (UMAP'21)*, no.25, pp.224-228, Utrecht, The Netherlands, June 2021.
- [9] J. W. Creswell and V. L. P. Clark, "Designing and Conducting Mixed Methods Research," SAGE Publications, Thousand Oaks, CA, 2007.
- [10] L. McInnes, J. Healy, and J. Melville, "UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction," *arXiv preprint arXiv:1802.03426 [stat.ML]*, Feb.2018.
- [11] H. L. O'Brien, P. Cairns, and M. Hall, "A Practical Approach to Measuring User Engagement with the Refined User Engagement Scale (UES) and New UES Short Form," *International Journal of Human-Computer Studies*, vol.112, pp.28-39, Apr.2018.
- [12] F. D. Davis, "Perceived usefulness, perceived ease of use, and user acceptance of information technology," *MIS Quarterly*, vol.13, no.3, pp.319-340, Sept.1989.
- [13] M. G. Armentano, I. Christensen, and S. Schiaffino, "Applying the Technology Acceptance Model to Evaluation of Recommender Systems," *Polibits*, Vol.51, no.1, pp.73-79, México, Mar./Jun.2015.
- [14] P. Pu and L. Chen, "A User-Centric Evaluation Framework of Recommender Systems," *Proc. ACM RecSys 2010 Workshop on User-Centric Evaluation of Recommender Systems and Their Interfaces (UCERSTI)*, pp.157-164, Barcelona, Spain, Sept.2010.
- [15] D. Lengyel, "Does the Netflix recommender system produce customer utility? - An analysis of the technology acceptance of the algorithmic-prediction-based Netflix recommender system and its drivers," *MSc dissertation, Universidade Católica Portuguesa and ESCP Europe Paris*, 79p., 2021.
- [16] C. L. Hsu and H. P. Lu, "Why do people play on-line games? An extended TAM with social influences and flow experience," *Information & Management*, vol.41, no.7, pp.853-868, Oct.2004.
- [17] I. Benbasat and W. Wang, "Trust in and adoption of online recommendation agents," *Journal of the Association for Information Systems*, vol.6, no.3, article 4, pp.1-30, Mar. 2005.
- [18] C. Flavián, M. Guinaliú, and R. Gurrea, "The role played by perceived usability, satisfaction and consumer trust on website loyalty," *Information & Management*, vol.43, no.1, pp.1-14, Jan. 2006.
- [19] S. Lee and J. Choi, "Enhancing user experience with conversational agent for movie recommendation: Effects of self-disclosure and reciprocity," *International Journal of Human-Computer Studies*, vol.103, pp.95-105, Mar.2017.
- [20] G. Norman, "Likert scales, levels of measurement and the "laws" of statistics," *Advances in Health Sciences Education*, vol.15, no.5, pp.625-632, 2010.