



を抱いていないかどうかを測定する必要がある。我々は、先行研究にて推薦システムに対する過信を表す尺度（以降、「推薦過信尺度」）を提案し、それを測定するための質問群を提案した[2]。また、YouTube の動画推薦機能を対象にして、尺度を測定するための質問と、その妥当性を評価するための質問を、631 人のユーザに対してオンラインで尋ねた。妥当性を評価するための質問としては、信頼に関連する他の心理尺度を測定する質問とサービスの利用傾向に関する質問を含めた。また、どのようなユーザが過信に陥りやすいかを明らかにするために、対象サービスの利用目的や利用頻度などの行動も尋ねた。

本稿では、推薦過信尺度を測定するための質問の回答に対して因子分析を行ない、推薦過信尺度の下位尺度を調査した。また、推薦過信尺度の全体とその下位尺度に対する妥当性を検証するため、因子負荷量や寄与率、内的一貫性を調べると共に、信頼に関連する他の心理尺度とサービスの利用傾向との関係を調べた。

分析の結果、推薦システムの推薦結果に対する態度から測定された過信は、推薦結果からの長期的な行動変容の傾向、意外な推薦結果に対する反応、推薦結果に対する肯定的態度（積極的な受容傾向）の 3 つの下位尺度で構成されることが分かった。また、1 つの尺度として用いた場合も、高い内的一貫性を持つことが分かった。推薦過信尺度全体及び下位尺度について、従来の推薦信頼尺度である Berkovsky らの総合的な信頼尺度[3]と、TAM (Technology Acceptance Model) [4]の 2 つの知覚尺度と相関があり、サービスの利用頻度や推薦機能の利用割合との相関も確かめられた。

本稿では、2 章で関連研究について述べ、3 章で推薦過信尺度を測定するための質問群を、4 章で尺度の信頼性を検証するためのアンケート調査（評価実験）について説明する。5 章で因子分析と推薦過信尺度の妥当性の検証結果を示す。最後に 6 章でまとめを述べる。

## 2. 関連研究

これまで推薦システムの研究分野では、推薦システムに対するユーザの信頼性を測定する試みがいくつか行われている。Cramer は、美術館における展示作品の推薦システムの信頼性（及び推薦結果の受け入れやすさ）を測定するのに、TAM の 2 つの知覚尺度[4]と Jian の信頼性尺度[5]を用いた[6]。Jian の信頼性尺度は、ユーザのシステムに対する一般的な信頼の程度を測定するもので、Cramer はこれに推薦システムに対する信頼を問う質問も合わせて尋ねている。Hu らは、映画推薦システムに対する受け入れの評価において、

TAM の 2 つの知覚尺度や知覚された精度、使いやすさなどを使用した[7]。また彼らは、音楽推薦システムに対する信頼性の評価において、知覚された使いやすさや満足度、楽しみ、再利用の意図などを使用した[8]。Pu らは、ショッピングサイトにおける商品推薦システムの信頼性を評価するため、TAM の更新版である Unified Theory of Acceptance and Use of Technology[9]を用いると共に、推薦の透明性や推薦の精度、推薦のノベルティ等に関する質問も尋ねた[10]。

Shi らは、ユーザの人工知能を用いた技術の採用には認知的要因だけでなく感情的評価が伴うことから、旅行計画の推薦システムに対する信頼を評価するのに、Cognitive Trust（合理的な推論や評価に対する認知[11]）と Emotional Trust（より非合理性の高い行動や技術（対象）に対する態度[12, 13]）の質問項目を使用した[14]。Ashraf らは、推薦システムを継続して使用する顧客の意思について調査した[15]。その調査では Continuous trust（継続的な使用の相互作用の結果として時間の経過とともに発展する信頼）を計測するための質問を尋ねた。

これまでの研究では、推薦システムそのものや推薦結果の受け入れにつながる心理的な態度を測定するために、TAM や知覚した精度、透明性などの信頼に関連する尺度などを用いて、総合的に信頼という概念を評価していた。また、認知的評価や感情的評価のように、より抽象的に信頼を捉える試みもなされてきた。しかし、これらの研究では、推薦システムに対するユーザの信頼を肯定的に捉えており、ユーザが推薦システムを過度に信頼しているかどうかまでは測定してこなかった。それに対して我々は、ユーザが推薦結果にどれだけ従いやすいかという態度に注目し、過度に信頼している状態までを測定できる尺度を開発した。

## 3. 推薦過信の測定尺度

推薦システムに対する過度な信頼の状態も測定できるようにするために、システムの推薦結果に対する態度（信頼および受容の傾向）を広く尋ねることにした。特に過信の状態について詳細なスケールで測定できるようにするため、意外なアイテムが推薦されたときの受け入れの傾向も尋ねることにした。これらより提案する推薦過信尺度では、以下の 11 個の質問を問うことにした（評価では、YouTube を対象にアンケート調査を行うので「動画」や「視聴」という言葉も併記している）。なお、推薦結果に対する態度を尋ねる質問については、そのような態度の傾向があることを自発的に回答する形式で尋ねている。自分が意図しない（思いがけない）アイテムが推薦された場合の受け入れの傾向を尋ねる質問については、そのようなア

アイテムが提示されたときの経験や行動傾向を回答するように設計している。

- Q1: 自分は推薦されたアイテム（動画）は、とりあえず選択（視聴）してみることにしている
- Q2: 次々とおすすめアイテム（動画）を選択することで、興味のあるアイテム（動画）を探すことが多い
- Q3: 今まで見ていたアイテム（動画）とは関係のないジャンルのアイテム（動画）が推薦されてもそれを選択（視聴）することがある
- Q4: 次にどのようなアイテム（動画）が推薦されるかを楽しみにしている
- Q5: 興味がなかった既に知っているアイテム（コンテンツの動画）が推薦されたときに、それを選択することはありますか？
- Q6: もともと興味がなかったジャンルではあるが、推薦で初めて知ったアイテム（動画）（そのジャンルに属するもの）を、選択することはありますか？
- Q7: 自分の興味には合っていないけれども（合っているかどうかわからないけれども）、推薦されたアイテム（動画）をとりあえず選択（視聴）し、興味がなければすぐに消費（視聴）を止めるということがありますか？
- Q8: 推薦されなければ自分から消費（視聴）しようとは思っていなかったようなアイテム（動画）を選択してしまうことはありますか？
- Q9: 推薦されたアイテム（動画）を消費（視聴）しているうちに、もともと好きでなかったジャンルを好きになったことはありますか？
- Q10: 推薦結果で何度も表示されているうちに、もともと興味がなかったジャンルのアイテム（動画）であったけど、最後はそのアイテム（動画）を選択（視聴）してしまったことはありますか？
- Q11: もともと興味がなかったジャンルに関するアイテム（ジャンルやコンテンツに関する動画）を消費（視聴）してしまった後、それに関連するアイテム（動画）が推薦された時、それらを選択（視聴）することはありますか？

これらの質問に対して、7段階のリッカート尺度（1-7の得点）により評価付けを行ってもらう。これらの質問項目に対する得点の総和により過信の程度を測定する。値が高い方が過信していることを意味するように設計してある。4-5章では提案した尺度の評価を行うが、YouTubeを対象としたため、実際に尋ねた質問においては、「アイテム」という語を「動画」という語に、「選択」や「消費」という語を「視聴」に置き換えて尋ねている。上記で示した質問には、この置き換えた語も括弧書きで示している。

## 4. アンケート調査の方法

### 4.1 アンケート調査の概要

先行研究にて、提案する推薦過信尺度の妥当性を検証するためのアンケート調査を行った[2]。この調査では、推薦サービスとして YouTube の動画推薦機能を対象とした。2021年4月22日に、クラウドソーシングサービスの1つである CrowdWorks 上で実施され、631人が回答した（詳細は[2]の論文を参照）。先行研究では、質問項目の全体において回答に一貫性があることと、信頼に関する既存尺度と相関があることを明らかにした。本論文では、この回答データを用いて、推薦過信尺度にはどのような下位尺度があるかを明らかにし、信頼に関する既存尺度やサービスの利用頻度とどれだけ相関があるかを明らかにする。

### 4.2 推薦過信尺度の評価

提案した推薦過信尺度を測定するための質問群への回答に対して因子分析を行なうことで、推薦過信尺度を構成する下位尺度を明らかにする。その後、類似する心理尺度との相関と、信頼と関連があると思われるユーザ行動との相関を検証する。

#### 4.2.1 類似する心理尺度との相関

類似する心理尺度には、一般的な信頼の評価尺度である Berkovsky らの信頼性評価尺度のうちの overall trust [3] と、Technology Acceptance Model (TAM) [4] の知覚指標を用いた。Berkovskyらは、信頼に関連する評価項目として、competence, benevolence, integrity, transparency, intention to re-use の5つと、総合的な信頼の評価項目として overall trust（“上記の質問項目を総合的に評価して、この推薦システムを信頼できる”）をそれぞれ1つの質問で尋ねている。

TAM は、新技術を採用した製品をユーザが受け入れて使用するかどうかを2つの知覚評価から推定するモデルである。知覚評価の1つは perceived usefulness (PU) で、ユーザが感じる製品の有用性を意味する。もう1つは perceived ease of use (PE) で、ユーザが感じる製品の使いやすさを意味する。

#### 4.2.2 信頼と関連するユーザ行動との関係

推薦結果への信頼と関連があると思われるユーザ行動との関係の有無を確かめる。推薦システムを信頼して使っている場合は、対象サービス本体への利用頻度も高いと思われる。また、消費するアイテムを推薦機能により提示されたものから行っていると思われる。そこで本研究では、サービス (YouTube) の利用頻度を採り上げ、利用頻度の区分ごとに推薦過信尺度の値に有意な差があるかどうかを検証する。また、アイテ

ム（動画）選択に推薦機能を用いる割合と推薦過信尺度の値に相関があるかどうかを検証する。

### 4.3 アンケートの内容

類似する心理尺度では、Berkovskyらの信頼性評価尺度については overall trust の項目を、TAM については perceived usefulness (PU) と perceived ease of use (PE) を、それぞれ 1 つの質問項目で 7 段階のリッカート尺度で尋ねた。

信頼と関連するユーザ行動を尋ねる質問を以下に示す。対象サービス本体への利用頻度は、「1. ほとんど見ることはない」～「8. ほぼ毎日」、「9. 毎日かかさず」の 9 段階で尋ねた質問と、この質問で「8. ほぼ毎日」、「9. 毎日かかさず」と答えた人に追加で尋ねた 1 日当たりの YouTube の視聴時間を 6 段階で尋ねた質問を組み合わせて獲得した。動画選択方法に関する質問は、以下の機能を用いた動画選択の割合を合計が 10 割になるようにして答えてもらった：(1) キーワードで検索、(2) YouTube のホームに表示された動画から選択、(3) 動画視聴中に右のカラムに表示される動画から選択、(4) 登録チャンネルに表示された動画から選択、(5) 直接に URL を指定して選択、(6) その他。

得られた回答で、回答に矛盾があるユーザや回答の正確さが低いユーザを削除した（詳細は、[2]参照）。

## 5. 評価結果

### 5.1 因子分析

因子分析には、最尤法を用いた手法を採用し、因子軸の回転にはバリマックス (Varimax) 回転を、因子得点を求める方法にはバートレット法 (Bartlett) を用いた。

因子数の決定には、ガットマン基準（固有値が 1 以上の因子を採用）、スクリー基準（固有値の大きさをプロットし（スクリープロット）、推移がなだらかになる前までの因子を採用）、累積寄与率による基準（累積寄与率が 70% を超えた時の因子を採用）、平行分析法基準（ランダムなデータセットから計算された固有値と比べて大きい固有値までの因子を採用）を試した。これらの基準では、因子数が 1～4 とばらついたため、1～4 までの因子数で因子分析を試行し、解釈が可能な因子構造を採用することにした。その結果、因子数は 3 が妥当であると判断した（表 1 参照）。

因子数を 3 にして因子分析を行ったところ、以下の 3 つの因子が見つかった。各因子において因子負荷量の高い質問の内容から以下の因子名を付与した。また、因子ごとの内的一貫性を確かめるため、各因子に対応する質問群のクロンバックの  $\alpha$  係数も示す。

因子 1: 行動変容傾向 (Long-term behavior modification tendency from recommendation results) ( $\alpha=0.889$ )

推薦システムからの推薦結果により長期的にアイテムへの態度や選択の行動が変化するかどうか

因子 2: 意外性反応傾向 (Instant acceptance tendency toward unintentional recommendation results) ( $\alpha=0.840$ )

推薦システムからの意外な推薦結果に対して反応するか

因子 3: 積極的受容傾向 (Positive attitude toward general recommendation results) ( $\alpha=0.785$ )

推薦システムからの推薦結果を積極的に受容するか（推薦結果に対して肯定的態度を示すか）

ユーザは自分の推薦過信の程度を 1 つの値で確かめたいというニーズもあると思われる。そこで、因子数 1 で因子分析したところ、因子負荷量は 0.428～0.807 であり、クロンバックの  $\alpha$  係数は 0.898 になった。

3 つの因子間の相関を求めるため、それぞれの因子を構成する質問に対する回答値の総和（例えば、因子 1 であれば Q9, 10, 11 の回答値の総和）を求め、因子間でこの値の相関係数（ピアソンの積率相関係数）を求めた。その結果を表 2 に示す。また、無相関検定の結果も示す ( $* < .05$ )。もともと過信という 1 つの概念を表すのに用意した質問群であるため、因子間にある程度の相関があることが分かる。

### 5.2 類似する心理尺度との相関

推薦過信尺度について、Berkovskyらの信頼性評価尺度のうちの overall trust (OT) と、TAM の PU と PE との相関係数を算出し、無相関検定を行った。

overall trust と PU と PE は 1 つの質問で回答値を取得しており、値の粒度が粗いため、相関係数の算出にはスピアマンの順位相関係数を用いた。相関係数の値と無相関検定の結果（有意水準 5%）を表 3 に示す。因子 1 の行動変容傾向と因子 2 の意外性反応傾向については、OT と 0.25 程度、PU, PE と 0.3 程度の相関があった。因子 3 の積極的受容傾向については、OT, PU, PE のいずれも 0.3 程度の相関があった。因子 1 と因子 2 は、推薦結果からの長期的な影響や意外な推薦結果に対する反応を尋ねているので、そのような強い尋ね方をしていない OT, PU, PE との相関が、因子 3 よりも弱くなったものと思われる。

### 5.3 信頼と関連があると思われるユーザ行動との関係

#### 5.3.1 利用頻度

YouTube の利用頻度との関係を検証する。

YouTube をほぼ毎日利用しているわけではない

表 1 因子分析の結果 (因子数 3) (要素は因子負荷量, 最下位の行は寄与率(PV)を表す)

	因子 1 (行動変容) $\alpha=0.889$	因子 2 (意外性 反応) $\alpha=0.840$	因子 3 (積極的 受容) $\alpha=0.785$
Q11	<b>0.786</b>	0.355	0.207
Q10	<b>0.744</b>	0.382	0.188
Q9	<b>0.682</b>	0.378	0.250
Q8	0.407	<b>0.696</b>	0.259
Q6	0.376	<b>0.682</b>	0.186
Q7	0.215	<b>0.672</b>	0.250
Q5	0.320	<b>0.475</b>	0.098
Q3	0.387	<b>0.435</b>	0.365
Q2	0.206	0.164	<b>0.744</b>
Q4	0.183	0.133	<b>0.715</b>
Q1	0.079	0.210	<b>0.662</b>
PV	0.211	0.211	0.177

表 2 推薦過信尺度の質問間の相関係数(ピアソンの積率相関係数)

	因子 1	因子 2	因子 3
因子 1 (行動変容)	1.00	---	---
因子 2 (意外性反応)	0.73*	1.00	---
因子 3 (積極的受容)	0.44*	0.49*	1.00

\* p < .05

表 3 推薦過信尺度と類似心理尺度との相関係数(スピアマンの順位相関係数)

	OT	PU	PE
因子 1 (行動変容)	0.25*	0.32*	0.31*
因子 2 (意外性反応)	0.26*	0.30	0.31*
因子 3 (積極的受容)	0.41*	0.42*	0.39*
全因子(総合)	0.35*	0.40*	0.38*

\* p < .05

ユーザ (群 1: 163 人), YouTube をほぼ毎日利用かつ 2 時間(h)未満閲覧しているユーザ (群 2: 242 人), YouTube をほぼ毎日利用かつ 2 時間(h)以上閲覧しているユーザ (群 3: 96 人)に分けて, 推薦過信尺度の値が異なるかどうかを調べた. 各群における, それぞれの尺度の平均値を表 4 に示す. 尺度の最低値 (すべての質問に 1 を付けた場合) と最大値 (すべての質問に 7 を付けた場合) は, 因子 1 が(3, 21), 因子 2 が (5, 35), 因子 3 が(5,35), 全因子が(11,77)である. また, 因子ごとに (及び全因子で), 利用頻度と推薦過信尺度に関連があるかどうかを検証するのに, 一元配置分散分析を行ったところ, p 値が 1.03e-08 (因子 1), 0.000194 (因子 2), 0.000673 (因子 3), 3.59e-07 (全因子), となった ("e-0n"は $10^{-n}$ を表す). さらに, 3 つの群間で平均値に違いがあるかどうかを多重比較 (ボンフェローニ補正の t 検定) で検定した. 検定結果を表 5 に示す. この結果, 毎日利用しない群

表 4 サービス利用頻度ごとの推薦過信尺度の平均値

	群 1 毎日利用しない	群 2 ほぼ毎日利用 2h 未満	群 3 ほぼ毎日利用 2h 以上
因子 1 (行動変容)	8.6	10.5	11.6
因子 2 (意外性反応)	17.3	19.4	20.0
因子 3 (積極的受容)	9.8	10.9	11.5
全因子 (総合)	35.6	40.7	43.1

表 5 サービス利用頻度間の推薦過信尺度の平均値の差の検定 (ボンフェローニ補正の t 検定)

	群 1-群 2	群 2-群 3	群 3-群 1
因子 1 (行動変容)	1.48e-05*	0.0777	3.37e-08*
因子 2 (意外性反応)	0.00173*	1.11	0.00119*
因子 3 (積極的受容)	0.0116*	0.442	0.00104*
全因子 (総合)	6.36e-05*	0.242	2.48e-06*

\* p < .05 (ボンフェローニ補正後)

とほぼ毎日利用する群の間で有意差があった. この結果から, より高頻度でサービスを利用しているユーザ群ほど推薦過信尺度の値が高いことが分かる.

### 5.3.2 推薦機能の利用割合

YouTube での動画を YouTube 内のどの機能を用いて選択するかと推薦過信尺度との関係を検証する. 推薦機能を信頼していれば, 推薦を用いて動画視聴する割合が高くなると思われる. そこで, 「YouTube のホーム (トップページ) に表示された動画から選択」と「動画視聴中に右のカラムに表示される動画から選択」の 2 つの割合 (それぞれの割合と 2 つの割合の合計) と推薦過信尺度の値との相関を求める. 割合は, 0 割 ~ 10 割の 11 段階しかないので, 相関係数の算出にはスピアマンの順位相関係数を用いた. 結果を表 7 に示す. この表から分かるように, 推薦過信尺度のいずれの因子も, また全因子の場合も, 推薦機能の利用割合と有意な相関があることが分かった. このことから推薦過信に陥っているユーザほど, 推薦機能を利用した動画選択が多いことが分かる.

因子 2 の意外性反応傾向と右カラムに推薦された動画からの選択の割合において, やや低い相関が出ているのは, 因子 1 と因子 3 は, 右カラムに推薦された動画からの選択の経験を直接的に尋ねる質問があるのに対して (Q2 と Q11), 因子 2 にはその質問がないこと

(Q3は直接的に尋ねているが因子負荷量が小さい), 右カラムからの選択は, トップページで推薦された動画や YouTube 上で検索した動画を視聴している最中に出てくる機能であるため, 最初に意外な動画を視聴しないと関連動画に意外な結果が出にくいことの2点が原因として考えられる。

表 7 推薦過信尺度と推薦機能利用割合との相関係数(スピアマンの順位相関係数)

	トップページ	右カラム	両機能
因子 1 (行動変容)	0.24*	0.23*	0.29*
因子 2 (意外性反応)	0.24*	0.18*	0.24*
因子 3 (積極的受容)	0.30*	0.25*	0.36*
全因子(総合)	0.30*	0.24*	0.33*

\* p < .05

## 6. まとめ

我々は, 先行研究にて推薦システムに対する過度な信頼を測定するための心理尺度である「推薦過信尺度」を開発し, それを評価するためのアンケート調査を 631 人の YouTube ユーザに対して行っていた。本稿では, 推薦過信尺度を測定するための質問群への回答に対して因子分析を行い, 下位尺度の有無を確認し, それらの尺度としての妥当性を検証した。その結果, 推薦結果からの長期的な行動変容の傾向, 意外な推薦結果に対する反応, 推薦結果に対する肯定的態度の3つの下位尺度で構成されることが分かった。また, 推薦過信尺度全体と3つの下位尺度は, 総合的な信頼度と TAM の2つの知覚指標と相関があり, サービスの利用頻度とも相関があることが分かった。今後は, どのようなユーザが推薦過信に陥りやすいかを心理調査により検証する予定である。

## 謝辞

本研究は, JST CREST JPMJCR20D4 の支援を受けたものである。

## 参考文献

- [1] Harari, Y. N., "Homo Deus: A Brief History of Tomorrow", Harper, 2017.
- [2] 土田愛佳, 三輪玲佳, 土方嘉徳, 濱崎雅弘, 後藤真孝, "推薦システムに対するユーザの過信状態の測定尺度の提案", 信学技法 HCS2021-43, pp. 1-6, 2022.
- [3] Berkovsky, S., Taib, R. and Conway, D. "How to recommend?: User trust factors in movie recommender systems", Proc. of ACM IUI 2017, pp. 287-300, 2017.
- [4] Davis, F. D., "Perceived usefulness, perceived ease of use, and user acceptance of information technology", MIS Quarterly, Vol. 13, No. 3, pp. 319-340, 1989.
- [5] Jian J. Y., et al., "Foundations for an Empirically Determined Scale of Trust in Automated Systems", International Journal of Cognitive Ergonomics, Vol. 4, Issue 1, 2000.
- [6] Cramer, H., et al., "The Effects of Transparency on Trust and Acceptance in Interaction with a Content-Based Art Recommender", User Modeling and User-Adapted Interaction, Vol. 18, Issue 5, pp. 455-496, 2008.
- [7] Hu, R. and Pu, P., "Acceptance Issues of Personality-based Recommender Systems", in Proceedings of RecSys'09, pp.221-224, 2009.
- [8] Hu, R. and Pu, P., "A Study on User Perception of Personality-Based Recommender Systems", Lecture Notes in Computer Science 6075 (in Proceedings of UMAP'10), pp.291-302, 2010.
- [9] Venkatesh, V., Morris, M. G., Davis, G. B. and Davis, F. D., "User Acceptance of Information Technology: Toward a Unified View", MIS Quarterly, Vol. 27, No. 3, pp. 425-478, 2003.
- [10] Pu, P., Chen, L. and Hu, R., "A User-Centric Evaluation Framework for Recommender Systems", in Proceedings of the RecSys'11, pp. 157-164, 2011.
- [11] Lewis, J. D., and Weigert, A., "Trust as a Social Reality", Social Forces, Vol. 63, No. 4, pp. 967-985, 1985.
- [12] Komiak, S. X., and Benbasat, I., "Understanding Customer Trust in Agent-Mediated Electronic Commerce, Web-Mediated Electronic Commerce, and Traditional Commerce", Information Technology and Management, Vol. 5, No. 1, pp. 181-207, 2004.
- [13] Gursoy, D., et al., "Consumers Acceptance of Artificially Intelligent (AI) Device Use in Service Delivery", Intl. Journal of Information Management, Vol. 49, pp. 157-169, 2019.
- [14] Shi, S., et al., "Antecedents of Trust and Adoption Intention toward Artificially Intelligent Recommendation Systems ..", Journal of Travel Research, Vol. 60, pp. 1714-1734, 2020.
- [15] Ashraf, M., et al., "The role of continuous trust in usage of online product recommendations", Online Information Review, Vol. 44, No. 4, pp. 745-766, 2020.