

# 自身の選択に対する固執性が他者の行動の予測に与える効果

## —行動実験と計算論モデリングによる検証—

大谷 佳名 (名古屋大学 大学院情報学研究科, otani.kanna@i.mbox.nagoya-u.ac.jp)

片平 健太郎 (名古屋大学 大学院情報学研究科, katahira.kentaro@b.mbox.nagoya-u.ac.jp)

Effects of choice perseverance on prediction of other's behaviour: Behavioral experiments and computational modeling

Kanna Otani (Graduate School of Informatics, Nagoya University, Japan)

Kentaro Katahira (Graduate School of Informatics, Nagoya University, Japan)

### Abstract

Predicting the behavior of each other is a fundamental process of social interaction. In order to ensure smooth communication, it is important not only to accurately predict the behavior of others but also to act in a way that is easily predicted by others. The choice perseverance, the tendency to repeat past choices, could play an important role in making one's behavior more predictable. To examine it, we conducted an experimental task in which participants predicted the choices of agents with different levels of perseverance. Here, we hypothesized that (1) the choices of those with high perseverance are more predictable than those with relatively low perseverance, and (2) participants with high perseverance can easily predict agents' choices with similarly high perseverance. To examine hypothesis (2), we quantified participant's perseverance by computational modeling. The experimental results showed that the accuracy of predictions in the block predicting choices of agents with high perseverance was significantly higher than in the block predicting choices of agents with relatively low perseverance ( $p < .01$ ). There is a significant correlation between the participant's perseverance estimated by the computational model and the prediction accuracy only when predicting choices of those with high perseverance ( $p < .01$ ). This result raises the possibility that people with similar behavioral tendencies are more likely to interact smoothly with each other. In conclusion, these results supported our hypotheses, and choice perseverance thus seems to be a key component in predicting others' behavior.

### Key words

perseverance, computational modeling, prediction of other's behavior, reinforcement learning, social interaction

### 1. はじめに

ヒトは円滑な社会生活を送る上で、他者の反応に応じて自身の行動を柔軟に調整することをしばしば求められる。例えば他者が怒りの表情を浮かべていれば、相手の神経を逆撫でしないように無難な行動をとることで軋轢を回避できる。このようなヒトの社会性は他者の行動を予測する能力によって支えられている。同時に、他者から予測されやすいように行動することも、他者に協力を促す上で重要な役割を持つと考えられている (Brennan & Clark, 1996; Vesper, van der Wel, Knoblich, & Sebanz, 2011)。他者から予測されやすい行動はどのような要素によって特徴づけられるのだろうか。一つの要素として、行動パターンを一貫させることが挙げられる。例えば、自動車の運転時に同じ車線を走り続けることには、後方のドライバーに対して自分(前方)の運転行動を予測させやすくするため、互いにとって安全な運転を促す効果がある。社会的相互作用における同じ行動を繰り返すことの効果を検討した研究として、複数人で共通の目標達成に向かって作業を行う共同行為(joint action)のパフォーマンスを向上させる (Vesper et al., 2011)、会話時に同じ単語を繰り返し参照することがコミュニケーションを円滑にする

(Brennan & Clark, 1996) といった報告がある。また、Wolf, Van Doorn, & Weissing (2011) は進化シミュレーションを用いて、一貫した選択行動は他者の予測を促すため、協力と競争の両側面において適応的意義を有すると指摘している。我々の日常は、時々刻々と変化する環境の中で過去の経験と現在の状況を照らし合わせ、価値の計算に基づいて最適な行動を選択する意思決定プロセスの連続であり、他者の行動を予測して自身の行動を調整する行為もその一例である。しかしながら、価値に基づく意思決定においても、一貫した選択行動が他者との交流を円滑にする効果は観察されるのかという疑問について明示的に実証した研究はない。一貫した選択行動が社会的な意思決定にどのような影響を与えるのか検討することは、我々が日々経験している他者との相互作用を支える認知機能を解明する上で重要であると考えられる。

近年では強化学習の枠組みから、同じ行動を繰り返す傾向(固執性: perseverance)が検討されている (Miller, Shenhav, & Ludvig, 2019; Wood & Rüniger, 2016)。本研究では固執性を行動の価値とは無関係に同じ選択を繰り返す傾向と定義する。例えば、視聴中のテレビ番組から有益な情報を得られないと気付いても、すぐには別の行動に移らずにテレビを見続けてしまう、という行為は固執的と言える。強化学習は選択の結果として得られる報酬に基づき行動の価値を計算し、最適な行動を逐次的に学習する過程である。しかし、実際的意思決定は価値の計算

のみによって行われるとは限らない。現在の選択（別の行動に移るか、テレビを見るか）は行動の価値（有益な情報は得られない）だけではなく、過去の選択（それでもテレビを見続ける）からも影響を受ける。最近の研究では、強化学習の枠組みを行動価値の変動と並行して過去の選択履歴も考慮するよう拡張することで、意思決定における固執性の影響を検討してきた（Katahira, 2018; 菅原・片平, 2019）。しかし、他者予測などの社会的行動における固執性の効果については十分に議論されていない。そこで、本研究は固執性と他者予測の関係について、強化学習の枠組みから検討することを目的とする。

これまで他者予測には、他者の選択や他者が受け取った報酬などの観察される情報に基づき行動パターンを推論する方略と、自身の行動パターンを他者に投影する方略の二つが考えられてきた。しかしながら、どちらの方略が優勢的に用いられているかは意見の分かれるところであった（Ames, 2004; Apperly, 2008）。一方で、二つの方略のどちらか片方ではなく、両方を組み合わせ用いているとの指摘もある（Hampton, Bossaerts, & O'Doherty, 2008; Suzuki, Harasawa, Ueno, Gardner, Ichinohe, Haruno, Cheng, & Nakahara, 2012）。Suzuki et al. (2012) は fMRI による脳機能計測と計算論モデリングと呼ばれる手法を用いて、ヒトが二つの方略を統合して他者予測を行っていることを示した。計算論モデリングとは、行動の背後にある心的過程を数理モデルで表現する手法である。実験から得られた行動データとモデルを照合することにより、モデルに含まれる構造の適合度（モデルがデータを説明する程度）や参加者の個人特性を定量的に評価することができる。本研究では Suzuki et al. (2012) にならって他者の選択を予測する行動実験と強化学習に基づく計算論モデル（以下、強化学習モデル）による分析を行った。本研究の第一の目的は、冒頭で述べたように同じ選択を繰り返すこと（固執性の高い行動）に自身の行動を予測させやすくする効果があるか否かを行動実験の結果から検討することである。そこで仮説 1 として、「固執性の比較的高いエージェント（実験条件における架空の他者）の選択は、それが比較的低いエージェントの選択よりも正確に予測しやすい」ことを検討した。そのため、実験で参加者は、固執性の高さが異なる二体のエージェントの選択をそれぞれ予測するよう求められた。参加者の予測とエージェントの選択が一致する割合を予測の精度（正確さ）の指標とし、固執性が比較的高いエージェントとそれが比較的低いエージェントの選択を予測する場合の予測の精度を比較することで仮説 1 を検討した。

次に、仮説 1 が支持された場合、固執性の高いエージェントの選択を予測しやすい理由としてどのようなことが考えられるだろうか。前述したように、他者予測には観察された情報に基づく推論と自己の投影との二つの方略が考えられているが、特に他者と自分との共通点が多く、類似性が高く認められる場合には「相手も自分と同じように考えるだろう」という投影が起りやすいという指摘がある（Ames, 2004）。自分と価値観や行動の傾向が似

ている他者とは馴染みやすく、良好な関係を築きやすいことは日常的に経験されるが、その理由として、投影による他者予測が正確に行われやすいことが考えられる。したがって、他者の固執性が高く、また自身の固執性も高い場合には、投影によって行動の予測が容易になる可能性がある。したがって、仮説 2 として「自身の固執性が高い人ほど同じように固執性の高いエージェントの選択を正しく予測できる」ことを検討した。固執性は個人によって高さが異なる特性であり、計算論モデリングを用いることで行動実験のデータに含まれるパラメータとして抽出することができる。行動実験から推定した参加者自身の固執性の高さ（パラメータ）と他者の選択に対する予測の精度との相関関係を個人ごとに比較することにより、検証を行った。なお、仮説が成り立つ前提として、本実験の参加者の行動に固執性の影響が認められることも確認した。その際に先行研究にならって固執性を組み込んだ強化学習モデルを作成し、それが固執性を含まないモデルよりもデータをより良く説明するか否かを検証した（Katahira, 2018; 菅原・片平, 2019）。

## 2. 方法

### 2.1 実験参加者

クラウドソーシングサービス CrowdWorks を利用する男女 200 名（男性 105 名、女性 95 名、平均年齢 39.32 歳、 $SD \pm 9.26$ ）が参加した。実験はすべてオンラインで実施した。参加者は Qualtrics 上で性別や年齢などの基本的な質問に答えた後、キーボード操作による行動課題を行った。行動課題は、画面の左右に表示される二つの選択肢のうちどちらかを、キーボードの「F」（左を選択）または「J」（右を選択）を押すことによる選択を繰り返すものである。なお、行動課題の作成と測定には Inquisit 5 を用いた。

参加者のうち、以下の基準に該当する者は分析対象から除外することとした。まず、行動課題の反応時間が早すぎる者は、選択肢を吟味せずにキーを押している可能性があるため、反応時間が 0.1 秒未満である試行が全体の 10 % 以上である者は除外した。次に、キー押し反応が制限時間である 5 秒を超過する試行が全体の 10 % 以上である者は、他の参加者と比較して試行数が減少するため除外した。最後に、次項で述べるように画面上に表示される選択肢の左右の位置は毎試行ランダムに入れ替わるため、選択肢とは無関係に同じキーを押し続けることは、課題の趣旨を理解した上での行動とは言い難い。また、本研究では検討する固執性は、キー押し反応に対する固執性ではなく、選択肢に対するものである。したがって、同じキーを押し続ける割合が全体の 20 % 以上である者は除外する。結果として 7 名を分析対象から除外し、193 名（男性 101 名、女性 92 名、平均年齢 40.01 歳、 $SD \pm 9.30$ ）のデータを分析に用いた。

### 2.2 実験手続き

実験デザインは Suzuki et al. (2012) にならい、ルールの

異なる2種類の行動課題から構成された(図1)。いずれの課題においても、選択肢として二種類のフラクタル図形(自己相似図形)を用いた。図1に示すように、選択肢の上には毎試行異なる得点が表示された。選択肢のうち一方は「当たり」となり、もう一方は「はずれ」となることとした。なお、選択肢の位置は毎試行ランダムに入れ替わった。

課題1(Self Task)は全90試行あり、参加者には課題全体でなるべく多くの得点を獲得するよう求めた。ルールは、「当たり」の選択肢を選んだ場合にその選択肢に割り当てられた得点を報酬として獲得できるものとした。「はずれ」の選択肢を選んだ場合、その回の獲得得点は0点とした。得点は選択肢(フラクタル図形)と同時に提示されるため、参加者は選択の前に得点の値を知ることになる。なお、「当たり」となる確率は選択肢ごとに異なるが、その値は参加者には知らされなかった。その確率は一方が0.75、他方が0.25に設定された。この確率は課題中で変わらないものとした。各選択肢の得点は全て整数とし、1点から99点までのいずれかが、両方を足して100点となる範囲で割り当てられた。なお、その得点は選択肢とは無関係に様分布からランダムに抽出された。また、得点の大きさが一方の選択肢に偏らないようにするため、同じ選択肢について50点以上の得点が3試行以上続かないという制限を設けた。参加者は5秒以内にどちらかを選択することが求められ、反応がなかった場合、その回の獲得得点は0点とした。選んだ選択肢の下に黄色のバーが0.5秒間表示された後、画面の中心に「当たり」の方の選択肢が1.5秒間表示された。続いて、獲得した得点と現時点での合計得点が1.5秒間表示された。

一方、課題2(Other Task)のルールは、Self Taskを行うエージェントの選択を一試行ごとに予測するという内容であった。参加者に対しては、過去に同一の実験に参加したエージェントのSelf Taskにおける選択を一試行ごとに予測するように教示した。参加者の予測とエージェ

ントの選択が一致すれば一律で50点の得点を獲得することができるが、予測が外れた場合の得点は0点とした。Self Taskと同様に、参加者が選んだ選択肢には黄色のバー、画面の中心には「当たり」の方の選択肢が表示されるほか、エージェントが選んだ選択肢には赤色の枠が1.5秒間表示された。その後、獲得得点と現時点での合計得点が1.5秒間表示された。

Other Taskのエージェントの選択として、強化学習モデルによって生成したシミュレーションデータを用いた。シミュレーションの詳細については後に述べる。Other Taskは2ブロック(各90試行)あり、選択への固執性が比較的低いエージェントを予測するブロック(Other Task 1)、固執性が比較的高いエージェントを予測するブロック(Other Task 2)から構成された。なお、Other Task 1とOther Task 2の順序は参加者内でカウンターバランスをとった。

### 2.3 モデル

各課題のデータの分析には、代表的な強化学習モデルの一つであるQ学習モデル(Qモデル)とリスク選好(受容・回避)を表すパラメータを追加したRiskモデル(RQモデル)、それぞれのモデルに対し固執性を表現するよう拡張したPerseveranceモデルとRisk with Perseveranceモデル(Q-Pモデル、RQ-Pモデル)の4つを用いた。すべてのモデルはSuzuki et al. (2012)のモデルに基づいて作成した。以下では、標準的なQ学習モデル(Qモデル)とPerseveranceモデル(Q-Pモデル)について簡潔に説明する。

まずSelf Taskのモデルについて述べる。Q学習モデルとは試行を繰り返すことで各選択肢の価値(Q値)を学習し、その差に基づいて確率的に選択を行うモデルである。各選択肢の得点の大きさが選択前に提示されるため、試行 $t$ での選択肢の価値 $Q(t)$ は、選択肢に割り当てられた得点 $R(t)$ 、「当たり」が出る確率の推定値 $p(t)$ を用いて次のように定義した(Behrens, Hunt, Woolrich, & Rushworth, 2008)。

$$Q(t) = p(t)R(t) \quad (1)$$

$p(t)$ は選択の結果(「当たり」か「はずれ」)をもとに更新されていくものとした。また、Qモデル、Q-PモデルにおいてQ値と $p(t)$ の更新式は同一のものとした。試行 $t$ において選ばれた選択肢について、「当たり」が出ると予測される確率 $p(t)$ は以下の更新式で表した。なお、選択の結果を $r(t)$ (「当たり」であれば1、「はずれ」であれば0)、予測と実際の結果との誤差を $\delta(t)$ とした。

$$\delta(t) = r(t) - p(t) \quad (2)$$

$$p(t+1) = p(t) + \alpha \delta(t) \quad (3)$$

その際、(3)式の学習率 $\alpha$ は $p(t)$ の更新の重み付けを決めるパラメータとし、これが大きいほど価値の学習は

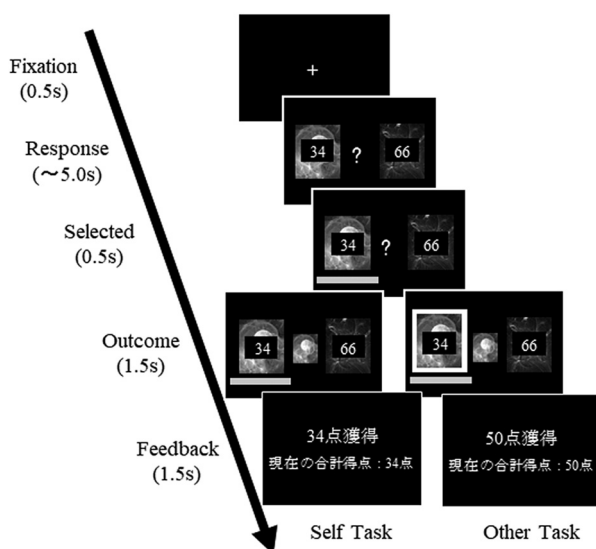


図1: 行動課題の一試行の流れ

直近の結果からの影響を受けやすくなる。また、 $p(t)$ の初期値は0とし、試行  $t$  において選ばれなかった選択肢の  $p(t)$  は更新しないものとした。

試行  $t$  で提示される選択肢を選択肢 A、選択肢 B とする。Q モデルにおいて、選択肢 A を選ぶ選択確率  $q_A(t)$  は、Q 値の差  $Q_A(t) - Q_B(t)$  を用いて以下のソフトマックス関数によって評価した。

$$q_A(t) = \frac{1}{1 + \exp(-\beta(Q_A(t) - Q_B(t)))} \quad (4)$$

逆温度  $\beta$  は、Q 値の差にどれだけ影響を受けて選択するかを表すパラメータとした。また、選択肢は2つであるため  $q_B(t) = 1 - q_A(t)$  が成り立つ。

一方、固執性の影響を含む Q-P モデルでは、選択時に過去の選択履歴の影響を考慮するようにソフトマックス関数を拡張した。その際、試行  $t$  までに各選択肢  $i$  がどれだけ選ばれてきたかと表す選択トレースと呼ばれる変数  $C_i(t)$  を定義し、試行  $t$  において選ばれた選択肢を  $a_t$  として、以下の更新式で更新した (Akaishi, Umeda, Nagase, & Sakai, 2014)。

$$C_i(t+1) = C_i(t) + \tau(I(a_t = i) - C_i(t)) \quad (5)$$

ここで、指示関数  $I(a_t = i)$  は試行  $t$  における選択が  $i$  であれば1、そうでなければ0をとるものとする。選択トレースは選択肢  $i$  が選ばれれば増加し、選ばれなければ減少する。 $\tau$  は選択履歴の減衰の速さを表すパラメータであり、1に近づくほど直前の選択の影響を受けやすく、0に近づくほどより遠い過去の選択を引きずりやすくなる。各選択肢の選択トレースを用いて、Q-P モデルにおける  $q_A(t)$  を以下のように設定した。

$$q_A(t) = \frac{1}{1 + \exp(-\beta(Q_A(t) - Q_B(t)) - \varphi(C_A(t) - C_B(t)))} \quad (6)$$

ここで、 $\varphi$  は過去の選択履歴をどれだけ選択に反映するかを示すパラメータであり、正であれば同じ選択を繰り返す傾向を、負であれば選択を切り替える傾向を表す。この関数において Q 値を含む項と選択トレースを含む項は独立した関係にあり、固執性は価値と無関係に同じ選択を繰り返す傾向であることを表している。

次に、エージェントの選択の予測を行う Other Task のモデルについて述べる。エージェントの選択の予測は以下のソフトマックス関数で計算した。Q-P モデルではエージェントの固執性を想定し、エージェントの選択履歴から計算した選択トレース  $C_o(t)$ 、選択肢 A の選択確率  $q_{oA}(t)$  を次のように設定した。また、固執性を含まない Q モデルでは選択トレースがないものとし、 $\varphi = 0$  とした。

$$q_{oA}(t) = \frac{1}{1 + \exp(-\beta(Q_{oA}(t) - Q_{oB}(t)) - \varphi(C_{oA}(t) - C_{oB}(t)))} \quad (7)$$

Self Task のモデルと同様に、Q 値は「当たり」が出ると予測される確率と得点の値を用いて  $Q_o(t) = p_o(t) R_o(t)$  と定義する。試行  $t$  で参加者が仮定する“エージェントが予測している「当たり」が出る確率”を  $p_o(t)$  として、以下の式で更新した。

$$p_o(t+1) = p_o(t) + \alpha_{RPE} \delta_o(t) + \alpha_{APE} \sigma_o(t) \quad (8)$$

$$\delta_o(t) = r_o(t) - p_o(t) \quad (9)$$

$$\sigma_o(t) = \frac{I(s_t = o_t) - p_o(t)}{R_o(t) p_o(t) (1 - p_o(t))} \quad (10)$$

ここで (9) 式の  $\delta_o(t)$  は、 $p_o(t)$  とエージェントの選択の結果  $r_o(t)$  との差分である“報酬に関する予測誤差” (Reward Prediction Error) とした。 $\delta_o(t)$  はエージェントが選んだ選択肢が「当たり」であれば正の値を取り、「はずれ」であれば負の値をとる。(10) 式の  $\sigma_o(t)$  は、エージェントが実際に選んだ選択肢  $o_t$  と参加者が予測した選択肢  $s_t$  との差分である“選択に関する予測誤差” (Action Prediction Error) とした。 $\sigma_o(t)$  は予測が一致すれば正の値をとる、予測が外れれば負の値をとる。(8)式において、 $\alpha_{RPE}$  は  $\delta_o(t)$  の、 $\alpha_{APE}$  は  $\sigma_o(t)$  の更新の重み付けをそれぞれ決める学習率とした。したがって  $p_o(t)$  は、報酬に関する予測誤差と選択に関する予測誤差の両方を用いて更新することとした (詳細は Suzuki et al., 2012 を参照のこと)。

## 2.4 パラメータ推定

ここでは、行動課題のデータからモデルのパラメータを推定する方法について簡潔に説明する。本研究では、各モデルのパラメータが個人ごとに異なると仮定し、個人特性を表す指標として用いることを目的としている。個人ごとのデータからパラメータ推定を行う場合、最尤推定などでは極端な値を持つ推定結果が得られやすいという問題がある (Katahira, 2016)。そこで、各パラメータの事前分布を想定する最大事後確率 (Maximum a posteriori; MAP) 推定を用いた。MAP 推定はベイズの定理によって与えられる事後分布の確率密度を最大化するパラメータ値を推定値として用いる方法である。本研究では、R の Rsolnp パッケージに含まれる solnp 関数を用いて非線形最適化によりパラメータを推定した。その際、全てのモデルにおいて学習率  $\alpha$  の事前分布はベータ分布 Beta (2,2) とし、 $0 \leq \alpha \leq 1$  の範囲をとるものとした。逆温度  $\beta$  は形状パラメータ 2、尺度パラメータ 0.03 のガンマ分布のもとで  $0 \leq \beta \leq 0.2$  とした。固執性を含むモデルでは、 $\tau$  はベータ分布 Beta (2,2) のもとで  $0 \leq \tau \leq 1$  とし、 $\varphi$  は平均 0、標準偏差 10 の正規分布を事前分布とした。

## 2.5 モデル比較

データを説明するモデルとしてどのモデルが適切かを判断するため、赤池情報量規準 (Akaike's information criterion, AIC) を用いて各モデルの適合度を比較した (Akaike, 1974)。その際に AIC の算出のため、各モデルについて最尤推定によって対数尤度を求めた。

## 2.6 シミュレーションデータの生成

ここでは Other Task のエージェントの選択として使用したデータの生成方法について述べる。固執性が比較的低いエージェントを予測するブロック (Other Task 1)、固執性が比較的高いエージェントを予測するブロック (Other Task 2) とともに、Q-P モデルを用いたシミュレーションによりデータを生成した。その際に、各パラメータについては同一の行動課題 (Self Task) を用いた予備実験のデータに基づき設定した。予備実験にはオンラインで 20 人 (男性 16 名、女性 4 名、平均年齢 41.8 歳、 $SD \pm 7.47$ ) が参加し、得られたデータから個人ごとに最尤推定によりパラメータ推定を行った。結果として、Other Task 1 のデータ生成には推定結果の中央値を用いて、学習率  $\alpha$  を 0.35、逆温度  $\beta$  を 0.027、 $\varphi$  を 0.076、 $\tau$  は 0.12 に設定した。Other Task 2 における  $\alpha$ 、 $\beta$ 、 $\tau$  も Other Task 1 と同一の値を用いたが、選択への固執性の高さを表すパラメータ  $\varphi$  については推定結果の第三四分位数をとり、1.85 に設定した。なお、全 90 試行の中で同じ選択を繰り返す試行数は、 $\varphi = 0.076$  の場合に 51 程度であり、 $\varphi = 1.85$  の場合に 69 程度となった。

## 2.7 統計解析

仮説 1 を検討するため、Other Task 1 と Other Task 2 における予測の精度について対応のある t 検定を行った。ここで各ブロックの予測の精度は、参加者の予測とエージェントの選択が一致した回数を試行数である 90 で割ることで算出した。また、モデルの適合度を統計的に比較するため、AIC に対して一要因の反復測定分散分析を行った。その際、球面性の仮定から逸脱した場合に Greenhouse-Geisser による自由度の調整を行った。その後の検定として、Shaffer の方法による多重比較を行った。次に、仮説 2 を検討するため Self Task における Q-P モデルのパラメータ  $\varphi$  と Other Task の予測の精度についてピアソンの積率相関係数を算出した。自身の固執性の高さ (Self Task から推定される  $\varphi$ ) と他者の固執性の高さ (Other Task の各ブロック) に交互作用があるか検討するため、各ブロックにおける相関係数について Thöni の方法による有意差検定を行なった (Thöni, 1977)。また、Self Task と Other Task それぞれにおける  $\varphi$  についても相関分析と相関係数の有意差検定を行った。

## 3. 結果

はじめに、Other Task における予測の精度はブロックによって差があるかを検討した。その結果、固執性が比較的高いエージェントを予測するブロック (Other Task 2)

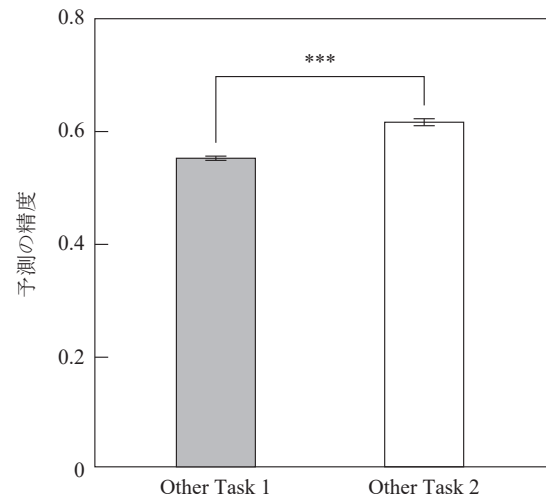


図 2: Other Task における予測の精度の平均値

注: アスタリスクは予測の精度の有意差を示す (\*\*\*)  $p < .01$ 。エラーバーは平均値の標準誤差を示す。

の予測の精度は固執性が比較的低いエージェントを予測するブロック (Other Task 1) と比較して有意に高かった ( $t(192) = 11.14, p < .01$ ; 図 2)。

また、Self Task のデータに対する各モデルの適合度 (AIC) を反復測定分散分析で比較した結果、モデル間の有意な差が認められた ( $F(1.07, 192) = 38.99, p < .01$ )。その後の検定により、固執性を含む Q-P モデルは他のモデルより有意に適合度が高く、最も良くデータを説明することが確認された ( $p < .01$ )。Other Task 1、Other Task 2 のデータについても同様の比較を行った結果、各モデルの適合度には有意な差があった ( $F(1.49, 192) = 21.91, p < .01$ ;  $F(1.79, 192) = 37.45, p < .01$ )。また、いずれのブロックにおいても固執性を考慮した Q-P モデルと RQ-P モデルは、他のモデルと比較して有意に適合度が高かった ( $p < .01$ )。Q-P モデルと RQ-P モデルの AIC に有意な差はなかったため (Other Task 1:  $p = .26$ ; Other Task 2:  $p = .27$ )、以降の分析ではよりシンプルな構造を持つ Q-P モデルを用いた。

各課題における参加者全員の行動データに対し、Q-P モデルに含まれるパラメータを推定した結果、推定結果の中央値は次のような結果となった。Self Task では  $\alpha$  が 0.43、 $\beta$  が 0.033、 $\varphi$  が 0.81、 $\tau$  が 0.32、Other Task 1 では  $\alpha_{RPE}$  が 0.31、 $\alpha_{APE}$  が 0.44、 $\beta$  が 0.046、 $\varphi$  が 0.57、 $\tau$  が 0.31、Other Task 2 では  $\alpha_{RPE}$  が 0.23、 $\alpha_{APE}$  が 0.39、 $\beta$  が 0.054、 $\varphi$  が 0.97、 $\tau$  が 0.27 となった。

次に、参加者の固執性の高さを定量化するため、過去の選択履歴をどれだけ選択に反映するかを示すパラメータ  $\varphi$  の値を Self Task、Other Task のデータから推定した。Self Task における  $\varphi$  の推定値と Other Task の予測の精度との間でピアソンの積率相関を求めた結果、Other Task 1 では  $r = .093$  ( $p = .19$ )、Other Task 2 では  $r = .35$  ( $p < .01$ ) であった (図 3)。よって、Other Task 2 にのみ参加者自身の固執性の高さと予測の精度は有意な相関が見られた。さらに、Other Task 2 の相関係数は Other Task 1 のそれと比較して

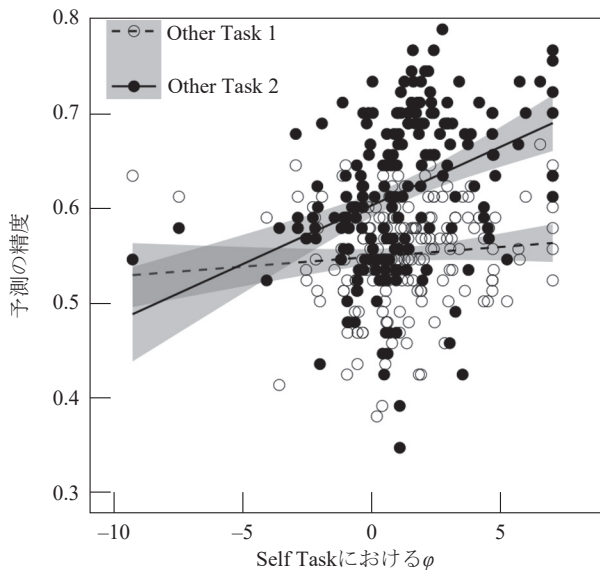


図3：Other Task の予測の精度と Self Task における参加者の固執性の高さ ( $\phi$ ) との相関図

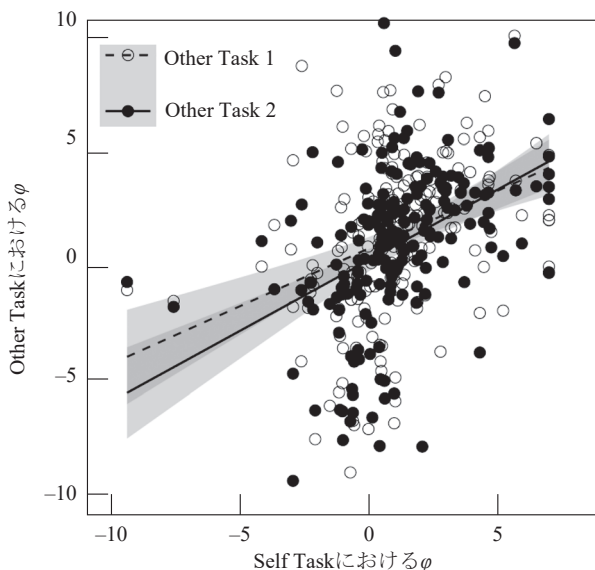


図4：Other Task と Self Task における固執性の高さ ( $\phi$ ) の相関図

有意に高かった ( $z = -2.6, p < .01$ )。

また、参加者自身の固執性とエージェントに想定する固執性との関係を検討するため、Self Task、Other Task それぞれの  $\phi$  の推定値について相関を求めた。結果、Other Task 1 では  $r = .35$  ( $p < .01$ )、Other Task 2 では  $r = .43$  ( $p < .01$ ) であり、いずれのブロックにおいても有意な正の相関が見られたが (図4)、ブロック間での相関係数の有意差は見られなかった ( $z = -0.92, p = .17$ )。

#### 4. 考察

本研究では、固執性の高い行動が他者予測を促進する可能性について行動実験と強化学習モデルを用いて検討

した。まず Self Task のモデル比較を行った結果、価値とは無関係に同じ選択を繰り返す固執性を組み込んだモデルが最もデータを良く説明していた。したがって、本研究の行動課題における参加者の振る舞いは、選択結果のみでなく過去の選択からも影響を受けることが確認されたため、仮説1と仮説2について分析を行った。

第一に、Other Task を行った参加者の予測の精度を比較したところ、固執性が比較的高いエージェントを予測する Other Task 2 の成績は、固執性が比較的低いエージェントを予測する Other Task 1 と比べて有意に高かった。よって、固執性の高いエージェントの選択はそれが比較的低いエージェントの選択よりも予測しやすいことが示された。この結果は仮説1を支持し、固執性には他者に対して自身の行動を予測させやすくする効果があることを示唆する。

第二に、参加者自身の固執性の強さを示すパラメータ  $\phi$  を Self Task のデータから推定し、Other Task での予測の精度と比較した。その結果、Other Task 2 においてのみ有意な正の相関が見られ、Other Task 1 と比較して強い相関が見られた。したがって、自身の固執性の高い人ほど同じように固執性の高い他者の選択を正しく予測できるという仮説2が支持された。さらに、Other Task のモデル比較を行った結果、他者の意思決定プロセスに固執性を想定するモデルが最もデータに適合していた。Self Task における参加者の  $\phi$  と、Other Task でエージェントに想定する  $\phi$  を比較した結果、Other Task 1 と Other Task 2 の両方において有意な正の相関が見られた。この結果から、自身の固執性の高い人ほど他者予測時に相手に対して高い固執性を仮定することが示された。

本研究の結果は、これまで自己完結的な行動特性として捉えられてきた固執性が社会的な機能を持つことを示唆するものである。仮説1が支持されたことから、固執性の高い行動は社会的文脈において他者からの正確な予測を促すことが実証された。しかしながら、本研究で用いた実験課題は一方向的に他者の選択を予測するものであり、囚人のジレンマゲームのように二者間の相互作用を検討するものではない。参加者には出来るだけ正確に予測を行うよう求めるに留まったため、より正確な予測がその後の行動にどのような影響を及ぼすかは不明である。したがって、Wolf et al. (2011) が進化シミュレーションで示したような、自身の固執性の高い人ほど他者からの協力を得やすいか否かという点については更なる検証が必要である。

また、本研究における仮説2の検証は、本稿の序章で述べた他者の行動を予測する際の二つの方略に焦点を当て、固執性の影響を検討したものであった。他者の行動の予測は、観察された情報に基づいて他者の行動パターンを学習する方略のみでなく、自身の行動パターンを他者に投影する方略によっても支えられている (Hampton et al., 2008; Suzuki et al., 2012)。実験の結果、自身の固執性の高い人ほど他者予測時に相手に対して高い固執性を仮定していること (図4)、自身の固執性の高い人ほど同じ

ように固執性の高い他者の選択を正しく予測できること(図3)が示された。これらの結果は、Ames (2004)の共通点の多い相手には自分との類似性を認知しやすくなるため、投影による予測が優勢的になる、という指摘と関連する。固執性の高い行動を予測しやすいことの理由としては、二者の行動パターンが類似しているため、投影によって正確な予測が行われやすくなることが考えられる。

本研究は、他者から予測されやすい行動を特徴づける要素の一つとして固執性の効果を検証した。二者間の相互作用に固執性の概念を導入することは、他者との交流に支障がある精神疾患における認知メカニズムの解明に寄与するだけでなく、人間が馴染みやすいロボットや自動運転車の開発などにも貢献する可能性がある。今後の研究として、固執性が他者予測を促進することは、向社会的な行動にどのように寄与するかを検討することが、固執性の真の社会的機能を明らかにする上で重要である。

#### 引用文献

- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19 (6), 716-723.
- Akaishi, R., Umeda, K., Nagase, A., & Sakai, K. (2014). Autonomous mechanism of internal choice estimate underlies decision inertia. *Neuron*, 81 (1), 195-206.
- Ames, D. R. (2004). Inside the mind reader's tool kit: Projection and stereotyping in mental state inference. *Journal of Personality and Social Psychology*, 87 (3), 340-353.
- Apperly, I. A. (2008). Beyond simulation-theory and theory-theory: Why social cognitive neuroscience should use its own concepts to study "theory of mind." *Cognition*, 107 (1), 266-283.
- Behrens, T. E. J., Hunt, L. T., Woolrich, M. W., & Rushworth, M. F. S. (2008). Associative learning of social value. *Nature*, 456 (7219), 245-249.
- Brennan, S. E. & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22 (6), 1482-1493.
- Hampton, A. N., Bossaerts, P., & O'Doherty, J. P. (2008). Neural correlates of mentalizing-related computations during strategic interactions in humans. *Proceedings of the National Academy of Sciences*, 105 (18), 6741-6746.
- Katahira, K. (2016). How hierarchical models improve point estimates of model parameters at the individual level. *Journal of Mathematical Psychology*, 73, 37-58.
- Katahira, K. (2018). The statistical structures of reinforcement learning with asymmetric value updates. *Journal of Mathematical Psychology*, 87, 31-45.
- Miller, K. J., Shenhav, A., & Ludvig, E. A. (2019). Habits without values. *Psychological Review*, 126 (2), 292-311.
- 菅原通代・片平健太郎 (2019). 強化学習における認知バ

イアスと固執性—選択行動を決めているのは過去の“選択の結果”か“選択そのもの”か?—。基礎心理学研究, 38 (1), 48-55.

- Suzuki, S., Harasawa, N., Ueno, K., Gardner, J. L., Ichinohe, N., Haruno, M., Cheng, K., & Nakahara, H. (2012). Learning to simulate others' decisions. *Neuron*, 74 (6), 1125-1137.
- Thöni, H. (1977). Testing the difference between two coefficients of correlation. *Biometrical Journal*, 19 (5), 355-359.
- Vesper, C., van der Wel, R. P. R. D., Knoblich, G., & Sebanz, N. (2011). Making oneself predictable: reduced temporal variability facilitates joint action coordination. *Experimental Brain Research*, 211 (3), 517-530.
- Wolf, M., Van Doorn, G. S., & Weissing, F. J. (2011). On the coevolution of social responsiveness and behavioural consistency. *Proceedings of the Royal Society B: Biological Sciences*, 278 (1704), 440-448.
- Wood, W. & Rünger, D. (2016). Psychology of habit. *Annual Review of Psychology*, 67 (1), 289-314.

(受稿：2020年5月5日 受理：2020年6月1日)