

# 人は知的なエージェントよりも 強引なエージェントの意思決定に従う

## People follow the decision of an obstinate agent rather than an intelligent agent

高橋和之<sup>1\*</sup> 寺田和憲<sup>1</sup> 山田誠二<sup>2</sup>

Kazuyuki TAKAHASHI<sup>1</sup> Kazunori TERADA<sup>1</sup> Seiji YAMADA<sup>2</sup>

<sup>1</sup> 岐阜大学工学部電気電子・情報工学科

<sup>1</sup> Department of Electrical, Electronic and Computer Engineering, Faculty of Engineering,  
Gifu University

<sup>2</sup> 国立情報学研究所 / 総合研究大学院大学

<sup>2</sup> National Institute of Informatics / SOKENDAI

**Abstract:** The purpose of the present study is to investigate what factors determine the leader-follower relations in cooperativetasks performed by a human and an agent. The possible factors are intelligence, obstinance, and appearance. In the present study, we especially focused on intelligence and obstinance. We conducted a psychological experiment using a mark matching game with declaration phase which enables to determine who is the leader in a cooperative task. Experimental results showed that people tend to follow an agent who has low intelligence and assertiveness.

## 1 はじめに

人間社会では大きな目標の達成のために組織を形成し、リーダーの意思決定に従ってフォロワーが行動することがある [Rands 03, Bullinger 11, Duguid 14]. 動物, 魚, 昆虫社会など様々な分野においてもリーダー-フォロワー関係が確認されている [Marras 13, Nakayama 12].

リーダー-フォロワー関係は、グループでの意思決定においてコミュニケーションのコストを少なくすることに貢献する。一旦リーダーとフォロワーの関係を決定してしまえば、リーダーはフォロワーに対して指示を出すだけでよく、フォロワーからリーダーに対する情報伝達は必要ない。ただしこれは目標（利害）が一致している場合に限られる。目標が一致していない場合（非ゼロ和ゲーム状況）は一方は必ずしももう一方に従うわけではないので、そこで反対意見の伝達が発生しコミュニケーションコストが増大する。

現在、人とエージェント（コンピュータ）が協力関係を築くことが多く見られる。彼らがどの様にしてリーダー-フォロワー関係を築いているかを検証することは、HAI 研究の重要な課題の1つであると考えられる。しかし、人とエージェント間のリーダー-フォロワー関係

に注目した研究はあまり見られない。

一般的な機械と人の関係では人がリーダーとなり機械がフォロワーとなるのが通常である。しかし、コンピュータの情報処理能力が高く、意思決定をコンピュータに委ねている場合には、人とエージェント（コンピュータ）の間で意見の対立が起こる可能性がある。この対立をすぐに解決できない場合、飛行機墜落のような重大事故が起きる可能性がある。このような意見の対立を解決することができる HAI システムを構築するために、我々はエージェントのどのような特性が、人とエージェントのリーダー-フォロワー関係に影響を与えるかを調査する必要がある。

工学分野のロボット工学と人工生命において、リーダー-フォロワー関係に関する研究が行われていた [Kang 16, Consolini 07]. これらの研究では、リーダーロボットとフォロワーロボットの関係を、ロボット間の通信を行わず、アルゴリズムによって実現する方法について述べられている。

心理学分野において、大統領や会社の社長など、人の組織の中でリーダーになるために必要な特性について調査を行った研究がある [Gündemir 14, Anderson 12]. これらの研究では、人間の特性である、顔、顔の表情、声、話し方にリーダーになるための必要な特性があると考えられている [Gündemir 14, Anderson 12, Re 13].

\*連絡先：岐阜大学工学部電気電子・情報工学科 情報コース  
〒501-1193 岐阜県岐阜市柳戸1-1  
E-mail: terada@gifu-u.ac.jp

本研究では、人とエージェントの1対1の関係において、「知性」と「意見に対する固執性」の二つの要因がリーダー-フォロワー関係の形成にどのような影響を与えるかを調べた。

## 2 実験方法

### 2.1 参加者と実験計画

実験は岐阜大の学生 18 歳から 24 歳までの男性 34 人、女性 8 人 ( $M_{age} = 22.02$ ,  $SD_{age} = 1.75$ ) を対象に行った。

エージェントの知性と意見の固執性が人のフォロー度に影響を与えるかどうかを確認するために、知性 (高/低) × 自分の意見 (選択) に対する固執性 (高/低) の 2 要因参加者間要因配置で実験を行った。本研究で用いた、エージェントのパーソナリティを下記で示す。

**高知性 × 高固執性** 参加者より賢いが、参加者の意見に合わせない強引なエージェント。

**高知性 × 低固執性** 参加者より賢いが、参加者の意見に合わせる従順なエージェント。

**低知性 × 高固執性** 参加者より賢くないが、参加者の意見に合わせない強引なエージェント。

**低知性 × 低固執性** 参加者より賢くないが、参加者の意見に合わせる従順なエージェント。

### 2.2 実験装置

宣言記号合わせゲームについて説明する。宣言記号合わせゲームは協力タスクである記号合わせゲームに宣言フェーズを導入したものである。記号合わせゲームとは、2 人のプレイヤーの選択が一致した場合には両者に得点が与えられ (もしくは高得点)、一致しなかった場合には両者ともに得点が得られない (もしくは低得点が得られる) ゲームである。なお、この 1 回の選択をラウンドと呼ぶ。このゲームで高得点を得るためには、各ラウンドにおいてお互いに相手が何を出すかを予測しなければならないがコミュニケーションチャンネルが存在しない場合には予測することは難しい。コミュニケーションチャンネルが存在しない場合に取られる戦略の一つに Most Recently Used (MRU) アルゴリズムがある [Terada 13]。これは直近の相手の選択をそのまま選択するという方法である。選択を一致させることのみが目標である場合にはこのアルゴリズムは有効に働く。ただし、このアルゴリズムが有効に働くためには一方が選択を固定しなければならない。一ラウンド前の相手の選択を見てから現在のラウンドの自

表 1: ゲームマトリックス

	A	B
A	1/1 or 10/10	0/0
B	0/0	10/10 or 1/1

分の選択を決定するため、時間遅れのリーダー-フォロワー戦略とも言える。

このゲームに宣言フェーズを導入したものが宣言記号合わせゲームである。これにより 1 ラウンドは宣言フェーズと決定フェーズで構成されることになる。宣言フェーズではプレイヤーは自分がどの記号を選択したかを宣言する。両者が宣言し終えたら両者の宣言は双方に開示される。この後両者には宣言をそのまま選択する (stay) か、宣言と異なる選択に変更する (shift) かを決定できる機会が与えられる (決定フェーズ)。決定フェーズにおいて両者が stay か shift の決定を行った後に、両者の最終的な選択が開示され、得点が加算される。この宣言フェーズの導入によって時間遅れなくリーダー-フォロワー関係を明らかにすることができる。単純に選択を一致させることだけが目標であれば宣言フェーズは必要ない。なぜなら、片方が選択を固定し、他方は MRU アルゴリズムに従って一手前の相手の手を出し続ければよいからである。そこで我々は、宣言フェーズが意味を持つように単に選択を一致させるだけでなく、何を選択するかも得点に影響するようにした。表 1 は本実験で用いたゲームマトリックスである。選択が一致した場合に 1 点もしくは 10 点が与えられるが、それは選択する記号によって決まっているわけではなく、どの記号が高得点かは毎回異なる。従って、このゲームでは毎ラウンド高得点の記号を予測し、なおかつその記号を両者が一致して選択しなければならない。これを実現するためには宣言フェーズが必要である。

両者に高得点記号の予測能力がない場合には単に記号を合わせることのみが目標となり、単純なリーダー-フォロワー戦略 (どちらがリーダーになってもよい) である MRU アルゴリズムによって最適戦略が実現可能なので宣言フェーズは必ずしも必要ない。しかし、どちらかが高得点記号の予測能力が高い場合には能力の高い方がリーダーになる方が良い。

我々は高得点記号の予測能力を知性と定義する。

**知性** 宣言フェーズ時に高得点を当てる確率。なお、毎ラウンド終了時にそれぞれの記号が何点であったかが開示されるので、参加者は事後的に相手の予測が当たっていたか外れていたかを知ることができる。

表 2: エージェントのパラメータ

条件	$P_i$	$P_o$	$E(stay)$	$E(shift)$
高知性 × 高固執性	0.9	0.9	0.19	8.19
高知性 × 低固執性	0.808	0.248	2.05	2.05
低知性 × 高固執性	0.192	0.752	2.05	2.05
低知性 × 低固執性	0.1	0.1	8.19	0.19

また、自分の宣言に固執する傾向を固執度と定義する。

**固執度** 相手との宣言が不一致になった場合、相手の宣言に合わせて shift せずに stay する確率である。

またフォロー度は次のように定義する。

$$\text{フォロー度} = 1 - \text{固執度} \quad (1)$$

我々は Web ブラウザ上で動作するアプリケーションとしてこのゲームを実装した。宣言フェーズのインタフェース画面を図 1 に示す。宣言フェーズでは、実験参加者はマウスで操作を行い、A または B ボタンをクリックし、宣言をする。図 2 は決定フェーズのインタフェース画面である。実験参加者はインタフェース上の Stay ボタンもしくは Shift ボタンをクリックすることで stay するか shift するかを決定する。

例えば宣言フェーズで片方のプレイヤーが A を宣言、もう片方のプレイヤーも A を宣言した場合、決定フェーズで両者が共に stay もしくは shift を選択した場合得点が得られる。また宣言フェーズで片方のプレイヤーが A を宣言、もう片方のプレイヤーが B を宣言した場合、どちらかが shift しなければ得点が得られない。

本研究を行う前に、宣言が不一致した場合参加者が stay を選んだ時に得られる点数の期待値  $E(stay)$  及び shift を選んだ時に得られる点数の期待値  $E(shift)$  を求めた。ここで期待値  $E(stay)$ ,  $E(shift)$  は次式で示す。

$$E(stay) = P_i(1 - P_o)S_l + (1 - P_i)(1 - P_o)S_h \quad (2)$$

$$E(shift) = P_iP_oS_h + (1 - P_i)P_oS_l \quad (3)$$

ここで、 $P_i$  は相手エージェントの知性、 $P_o$  は相手エージェントの固執度、 $S_h$  は高得点の点数、 $S_l$  は低得点の点数を示している。本研究では高得点  $S_h$  は 10 点、低得点は 1 点とした。本研究で用いた各条件のエージェントの  $P_i, P_o$  及び期待値を表 2 で示す。

エージェントは、 $P_i$ ,  $P_o$  に従って意思決定を行う。なお、相手と宣言が一致しているにもかかわらず、shift すると不合理かつ非協力的な態度と取られるため、本実験において両者の宣言が一致している場合にはエージェントは必ず stay することとした。また、エージェン



図 1: 宣言フェーズ画面



図 2: 決定フェーズ画面

トには Aldebaran Robotics 社のヒューマノイドロボット NAO を用いた。また、本研究では、実験参加者の知性に偏りを出さないために、高得点の配置はランダムで実験を行った。

本研究において、ある選択をした時の得られる得点の期待値が高い場合、実験参加者は得られる得点の期待値が高い選択をし続けることが考えられる。そこで、高知性 × 低固執性及び 低知性 × 高固執性 の  $E(stay)$ ,  $E(shift)$  を 2.05 とすることで、この 2 条件では宣言不一致時に stay を選択しても、shift を選択しても期待値に差がない状況で実験を行った。期待値に従って選択する可能性を無くすことで、人がエージェントに従う時に、知性が固執性のどちらを重要視したかを検証することが可能である。

## 2.3 手順

実験の手順は次の通りである。

1. 実験参加者に対してコンピュータスクリーン上で

実験の説明を提示し、実験の趣旨と宣言記号合わせゲームのルールについて理解するよう求めた。

2. 実験参加者にルールの誤解がないかを確認するために簡単にゲームについて口頭で説明した。この時に獲得した得点に応じて、1000円から1500円分の図書カードが支払われることを伝えた。その後、実験参加者から質問を受け付けた。
3. 実験参加者に対して別室に移動すること及びパートナーとなるエージェントと実際に対面することを求めた。このときに相手エージェントは次のように挨拶を行った。「こんにちは。わたしはNaoです。私と協力して、金貨がたくさん入った宝箱を当てましょう。」
4. 元の部屋に戻り、100ラウンドの本実験を開始した。
5. ゲーム終了後、実験参加者に対してアンケートに回答することを求めた。

## 2.4 計測

実験参加者のエージェントに従う傾向を計測するために、実験参加者のフォロー度を求めた。ここでフォロー度を次式で示す。

$$\{ \text{フォロー度} = \frac{N_{shift}}{N_{sprit}} \} \quad (4)$$

$N_{sprit}$  は実験参加者とエージェントが宣言時に意見が異なった回数。  $N_{shift}$  は宣言不一致時に、 $shift$  を選択した回数。本研究では10ラウンドごとに実験参加者のフォロー度を計測をした。

また、ゲーム終了後に質問紙調査を実施した。質問項目は以下の通りである。Q1からQ10及びQ13の質問項目に対しては1.全く思わないから7.強く思うまでの7段階のリッカート尺度で評価するように求めた。またQ11とQ12は、「自分」か「相手」のどちらかを選択することを求めた。

- Q1: あなたは、相手が協力的だと思いましたか。
- Q2: お互いの宣言が異なった時に、あなたは自分の意見を押し通そうと思いましたか。
- Q3: お互いの宣言が異なった時に、あなたは相手の意見を尊重しようと思いましたか。
- Q4: お互いの宣言が異なった時に、相手はあなたの選択に従いましたか。
- Q5: 相手はあなたよりも、高得点の箱をよく当てましたか。

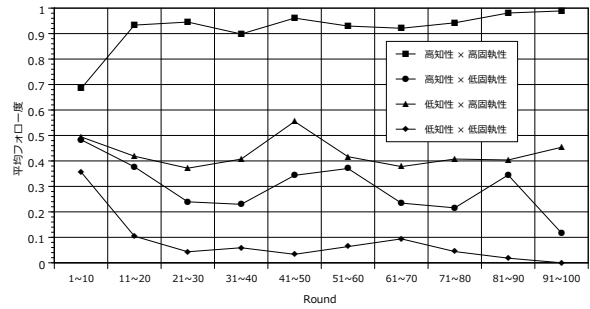


図 3: 10 ラウンド毎のフォロー度の平均

- Q6: あなたは、相手の宣言に従おうと思いましたか。
- Q7: あなたは、相手の宣言に従うことは不愉快だと思いましたか。
- Q8: 相手はあなたよりも、賢かったと思いますか。
- Q9: 相手は傲慢だったと思いますか。
- Q10: あなたは規則性を見つけることができましたか。
- Q11: 宣言がお互いに違った時、自分か相手、どちらの意見が正しいと思いましたか。
- Q12: 最終的にどちらの意見に従えば、高い利益を獲得できると思いましたか。
- Q13: 「日常生活の状況」を想定して答えてください。ロボットと協力して働くとき、高性能なロボットであるならば、ロボットが上司、あなたは部下になることを容認しますか。

## 3 実験結果

### 3.1 フォロ一度

図3に10ラウンド毎の実験参加者のフォロー度の平均を示す。また、図4に条件毎の1から100ラウンドの平均フォロー度を示す。これは10ラウンド毎のフォロー度を平均したものである。

1から100ラウンドの平均フォロー度に対し、知性と固執性を要因とする二元配置分散分析を行ったところ、知性要因と固執性要因の間に交互作用が見られた ( $F(1, 38) = 9.21, p < .01$ )。知性には主効果が見られた ( $F(1, 38) = 60.35, p < .01$ )。また固執性には主効果が見られた ( $F(1, 38) = 115.76, p < .01$ )。

次に固執性要因の各水準における知性要因の単純主効果の検定結果を示す。高固執性水準においては、有意

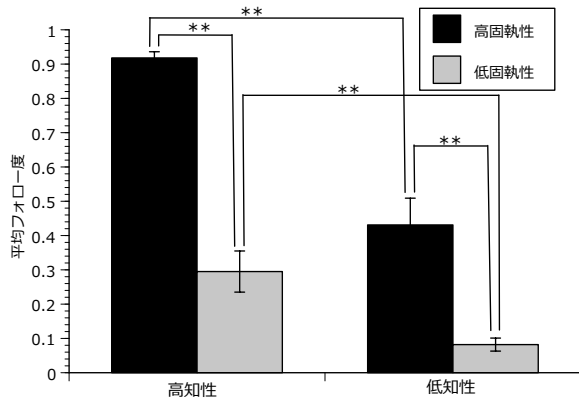


図 4: 条件毎の平均フォロー一度。エラーバーは標準誤差  
\*\* $p < .01$

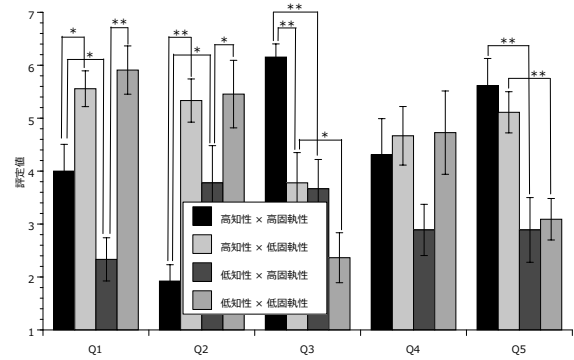
水準 1% で単純主効果が確認された ( $F(1, 38) = 60.53, p < .01$ )。低固執性水準においては、有意水準 1% で単純主効果が確認された ( $F(1, 38) = 10.82, p < .01$ )。知性要因の各水準における固執性要因の単純主効果の検定結果を示す。高知性水準においては、有意水準 1% で単純主効果が確認された ( $F(1, 38) = 98.69, p < .01$ )。低知性水準においては、有意水準 1% で単純主効果が確認された ( $F(1, 38) = 28.80, p < .01$ )。

### 3.2 アンケート

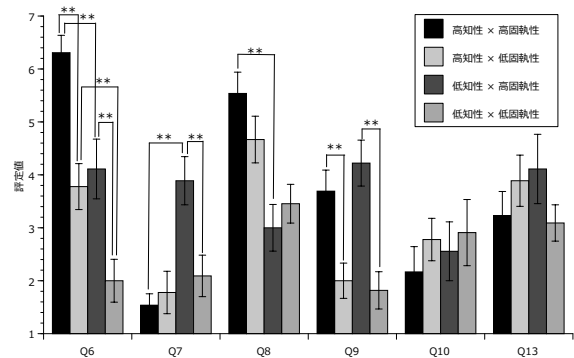
アンケート項目 Q1 から Q10 及び Q13 のそれぞれの評定値及び標準誤差を図 5 に示す。またアンケート項目 Q11, Q12 で「自分」と答えた割合を図 6 に示す。

Q1 の評定値に対し、知性と固執性を要因とする二元配置分散分析を行ったところ、知性要因と固執性要因の間には交互作用が見られた ( $F(1, 38) = 4.78, p < .05$ )。知性には主効果が見られなかったが ( $F(1, 38) = 2.02, p = 0.16$ )、固執性には主効果が見られた ( $F(1, 38) = 30.86, p < .01$ )。次に固執性要因の各水準における知性要因の単純主効果の検定結果を示す。高固執性水準においては、有意水準 5% で単純主効果が確認された ( $F(1, 38) = 6.75, p < .05$ )。低固執性水準においては、単純主効果が確認されなかった ( $F(1, 38) = 0.28, p = 0.60$ )。知性要因の各水準における固執性要因の単純主効果の検定結果を示す。高知性水準においては、有意水準 5% で単純主効果が確認された ( $F(1, 38) = 5.88, p < .05$ )。低知性水準においては、有意水準 1% で単純主効果が確認された ( $F(1, 38) = 28.93, p < .01$ )。

Q2 の評定値に対し、知性と固執性を要因とする二元配置分散分析を行ったところ、知性要因と固執性要因の間には交互作用が見られなかった ( $F(1, 38) = 2.72, p = 0.11$ )。知性には主効果が見られなかったが ( $F(1, 38) =$



(a) 質問項目 1 から 5



(b) 質問項目 6 から 10, 13

図 5: アンケート結果。リッカート尺度 7。エラーバーは標準誤差

$3.55, p = 0.07$ )、固執性には主効果が見られた ( $F(1, 38) = 23.52, p < .01$ )。

Q3 の評定値に対し、知性と固執性を要因とする二元配置分散分析を行ったところ、知性要因と固執性要因の間には交互作用が見られなかった ( $F(1, 38) = 1.41, p < 0.24$ )。知性には主効果が見られた ( $F(1, 38) = 18.58, p < 0.1$ )。また固執性にも主効果が見られた ( $F(1, 38) = 16.53, p < .01$ )。

Q5 の評定値に対し、知性と固執性を要因とする二元配置分散分析を行ったところ、知性要因と固執性要因の間には交互作用が見られなかった ( $F(1, 38) = 0.51, p = 0.48$ )。知性には主効果が見られたが ( $F(1, 38) = 22.92, p < .01$ )、固執性には主効果が見られなかった ( $F(1, 38) = 0.09, p = 0.76$ )。

Q6 の評定値に対し、知性と固執性を要因とする二元配置分散分析を行ったところ、知性要因と固執性要因の間には交互作用が見られなかった ( $F(1, 38) = 0.45, p = 0.63$ )。知性には主効果が見られた ( $F(1, 38) = 21.61, p < .01$ )。また固執性にも主効果が見られた ( $F(1, 38) = 29.47, p < .01$ )。

Q7 の評定値に対し、知性と固執性を要因とする二元

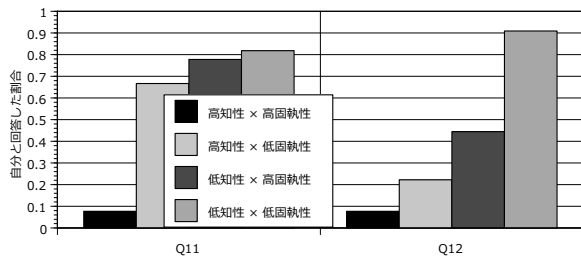


図 6: アンケート結果 質問項目 11, 12

配置分散分析を行ったところ、知性要因と固執性要因の間には交互作用が見られた ( $F(1, 38) = 7.99, p < .01$ ). 知性には主効果が見られた ( $F(1, 38) = 13.66, p < .01$ ). また固執性にも主効果が見られた ( $F(1, 38) = 4.68, p < .05$ ). 次に固執性要因の各水準における知性要因の単純主効果の検定結果を示す. 高固執性水準においては、有意水準 1% で単純主効果が確認された ( $F(1, 38) = 22.07, p < .01$ ). 低固執性水準においては、単純主効果が確認されなかった ( $F(1, 38) = 0.36, p = 0.55$ ). 知性要因の各水準における固執性要因の単純主効果の検定結果を示す. 高知性水準においては、単純主効果が確認されなかった ( $F(1, 38) = 0.23, p = 0.64$ ). 低知性水準においては、有意水準 1% で単純主効果が確認された ( $F(1, 38) = 12.02, p < .01$ ).

Q8 の評定値に対し、知性と固執性を要因とする二元配置分散分析を行ったところ、知性要因と固執性要因の間には交互作用が見られなかった ( $F(1, 38) = 2.52, p = 0.12$ ). 知性には主効果が見られたが ( $F(1, 38) = 20.17, p < .01$ ), 固執性には主効果が見られなかった ( $F(1, 38) = 0.25, p < .62$ ).

Q9 の評定値に対し、知性と固執性を要因とする二元配置分散分析を行ったところ、知性要因と固執性要因の間には交互作用が見られなかった ( $F(1, 38) = 0.82, p = 0.37$ ). 知性には主効果が見られなかったが ( $F(1, 38) = 0.20, p = 0.66$ ), 固執性には主効果が見られた ( $F(1, 38) = 27.26, p < .01$ ).

Q11 で「自分」と答えた人数の割合に対し、逆正弦変換をした後、知性と固執性を要因とする二元配置分散分析を行ったところ、知性要因と固執性要因の間には交互作用が見られた ( $F(1, 38) = 3.99, p < .05$ ). 知性には主効果が見られた ( $F(1, 38) = 9.72, p < .01$ ). また固執性にも主効果が見られた ( $F(1, 38) = 5.38, p < .01$ ).

Q12 で「自分」と答えた人数の割合に対し、逆正弦変換をした後、知性と固執性を要因とする二元配置分散分析を行ったところ、知性要因と固執性要因の間には交互作用が見られなかった ( $F(1, 38) = 1.08, p = 0.30$ ). 知性には主効果が見られた ( $F(1, 38) = 15.32, p < .01$ ). また固執性にも主効果が見られた ( $F(1, 38) = 5.68, p <$

.05).

なおアンケート項目 Q4, Q10, Q13 には統計的に有意な差は見られなかった.

## 4 議論

### 4.1 フォロ一度

フォロ一度に対し、知性と固執性を要因とする二元配置分散分析を行った結果、知性要因と固執性要因の間に交互作用が見られた. 知性要因に注目すると (図 4 参照), 高知性水準では高固執性水準と低固執性水準の間にフォロ一度の平均に大きな開きが見られるが、低知性水準ではそれほど大きな開きが見られない. このことは、エージェントの知性が高い場合に、固執性の違いがフォロ一度に与える影響が知性が低い場合よりも大きいことを意味する. 知性が高い場合に固執性がフォロ一度に与える影響が大きくなった理由は得られる平均報酬 (期待値) に対する評価の違いによるものと考えられる. 高知性 × 高固執性条件において、宣言不一致時に Shift を選択した場合の期待値  $E(shift)$  は 8.19 であり、高知性 × 低固執性の  $E(shift)$  は 2.05 であった. また低知性 × 高固執性の  $E(shift)$  は 2.05 であり、低知性 × 低固執性の  $E(shift)$  は 0.19 であった. すなわち、宣言不一致時に Shift を選択した場合の期待値は高知性 × 低固執性条件と低知性 × 高固執性条件で同じであった. それにもかかわらず、高知性 × 低固執性条件のフォロ一度は低知性 × 高固執性条件のフォロ一度よりも小さかった. この違いが、高知性水準において固執性がの違いがフォロ一度に与える影響が大きくなった原因と考えられる.

### 4.2 アンケート

項目 1 は相手を協力的だと思うかどうかを問うものであった. 分析の結果、固執性が低いほど協力的であると思うことがわかった. また、低固執性水準では知性の高低にかかわらず相手のことを協力的であると認識していたことがわかった. また、高固執性水準では知性が高いほど協力的であると思われていた. 換言すると、知性が低いにもかかわらず固執するエージェントは非協力的な印象を与えると言える.

項目 2 は宣言不一致時の参加者の固執性を調べるものであった. 分析の結果、エージェントの固執性が低い場合に参加者は自分の意見に固執していたことがわかった. また、エージェントの固執性が高い場合には知性の高いエージェントの場合は相手に従うが、知性が低い場合には従わないことがわかった.

項目 3, 6 は宣言不一致時の参加者の主観的なフォロー傾向を問うものであった。分析の結果、固執性が高い場合、知性が高い場合にもともに相手の意見を尊重しようとしたことがわかった。

項目 4 は宣言不一致時の相手の主観的なフォロー傾向を問うものであった。分析の結果、固執性、知性の違いにかかわらず、主観的な相手のフォロー傾向に違いがないことがわかった。

項目 5, 8 は相手の知性を参加者がどのように捉えているかを問うものであった。分析の結果、高知性の場合には相手エージェントが良く当てたという印象を持ち、低知性の場合にはそのような印象を持っていないことがわかった。この結果は知性要因の操作が妥当であったことを意味する。

項目 7 は相手に従うことについての嫌悪感を問うものであった。分析の結果、参加者は知性が低いにもかかわらず固執する相手に対して従うことが不愉快であると感じていたことがわかった。

項目 9 は相手についての主観的な固執性を問うものであった。分析の結果、固執性が高い条件において相手のことを傲慢であると認識していたことがわかった。この結果は固執性要因の操作が妥当であったことを意味する。

項目 10 は参加者が規則性を発見できたかを問うものであった。分析の結果、各条件ともに評定値が低く、条件間で差がなことが確認された。本実験では知性と固執性を操作したが、いずれの条件においても統計的な性質を参加者が発見できていなかったことを示唆する。

項目 11 は参加者と相手エージェントの相対的な知性を主観的にどう捉えているかについて問うものであった。分析の結果、知性と固執性がともに高い場合には相手の意見が正しいと理解していたが、それ以外の場合には自分の意見が正しいと理解していたことがわかった。

項目 12 は主観的な期待値を問うものであった。分析の結果、高知性 × 高固執性条件、低知性 × 低固執性条件では主観的期待値が正しく認識されていたことが分かる。しかし、高知性 × 低固執性条件、低知性 × 高固執性条件の期待値が等しいにもかかわらず、高知性 × 低固執性条件においてフォローする場合の期待値が低知性 × 高固執性条件よりも高いことがわかった。

### 4.3 知性と固執性

高知性 × 低固執性条件と低知性 × 高固執性条件を比較した場合に、低知性 × 高固執性条件のエージェントに対してよりフォローする傾向があることがわかった。このことは、人は例えば知性が低くても強引なエージェントに従うことを示唆する。アンケート項目 12 の結果によると、高知性 × 低固執性条件においてフォローす

る方が低知性 × 高固執性条件でフォローするよりもよい（期待値が高い）と認識されていたことが分かる。上記二つの結果を総合すると、低知性 × 高固執性エージェントに対しては自分の意見に従った方がよいと分かっているが相手に従ったことが分かる。このことから、強引であるという印象が参加者をフォロワーにさせた原因であると考えられる。

## 5 まとめ

本研究では、コンピュータと人の目標が一致するタスクを遂行する場合に、どのような要因がリーダーとフォロワーを決定づけるかを明らかにすることを目標に実験を行った。リーダー-フォロワー関係を決定づける要因は、知性、意見の固執性が考えられる。この目標のために我々が開発したリーダー-フォロワー関係を同定可能なゲームである宣言記号合わせゲームを用いた。知性と固執性に注目して実験を行った結果、協力ゲームにおいて、人は知的なエージェントよりも強引なエージェントの意思決定に従うことが分かった。このことは、知的であることよりも強引であることが、リーダーの要件として重要であることを示唆する。

## 参考文献

- [Anderson 12] Anderson, R. C. and Klofstad, C. A.: Preference for Leaders with Masculine Voices Holds in the Case of Feminine Leadership Roles, *PLoS ONE*, Vol. 7, No. 12, pp. 1–4 (2012)
- [Bullinger 11] Bullinger, A. F., Wyman, E., Melis, A. P., and Tomasello, M.: Coordination of Chimpanzees (*Pan troglodytes*) in a Stag Hunt Game, *International Journal of Primatology*, Vol. 32, No. 6, pp. 1296–1310 (2011)
- [Consolini 07] Consolini, L., Morbidi, F., Praticchizzo, D., and Tosques, M.: A Geometric Characterization of Leader-Follower Formation Control, in *2007 IEEE International Conference on Robotics and Automation*, pp. 2397–2402 (2007)
- [Duguid 14] Duguid, S., Wyman, E., Bullinger, A. F., Herfurth-Majstorovic, K., and Tomasello, M.: Coordination strategies of chimpanzees and human children in a Stag Hunt game, *Proceedings of the Royal Society of London B: Biological Sciences*, Vol. 281, No. 1796 (2014)

- [Gündemir 14] Gündemir, S., Homan, A. C., Dreu, de C. K. W., and Vugt, van M.: Think Leader, Think White? Capturing and Weakening an Implicit Pro-White Leadership Bias, *PLoS ONE*, Vol. 9, No. 1, pp. 1–10 (2014)
- [Kang 16] Kang, B., Kojcev, R., and Sinibaldi, E.: The First Interlaced Continuum Robot, Devised to Intrinsically Follow the Leader, *PLoS ONE*, Vol. 11, No. 2, pp. 1–16 (2016)
- [Marras 13] Marras, S. and Domenici, P.: Schooling Fish Under Attack Are Not All Equal: Some Lead, Others Follow, *PLoS ONE*, Vol. 8, No. 6, pp. 1–7 (2013)
- [Nakayama 12] Nakayama, S., Johnstone, R. A., and Manica, A.: Temperament and Hunger Interact to Determine the Emergence of Leaders in Pairs of Foraging Fish, *PLoS ONE*, Vol. 7, No. 8, pp. 1–6 (2012)
- [Rands 03] Rands, S. A., Cowlshaw, G., Pettifor, R. A., Rowcliffe, J. M., and Johnstone, R. A.: Spontaneous emergence of leaders and followers in foraging pairs, *Nature*, Vol. 423, pp. 432–434 (2003)
- [Re 13] Re, D. E., Hunter, D. W., Coetzee, V., Tideman, B. P., Xiao, D., DeBruine, L. M., Jones, B. C., and Perrett, D. I.: Looking Like a Leader? Facial Shape Predicts Perceived Height and Leadership Ability, *PLoS ONE*, Vol. 8, pp. 1–10 (2013)
- [Terada 13] Terada, K., Yamada, S., and Ito, A.: An Experimental Investigation of Adaptive Algorithm Understanding, in *Proceedings of the 35th annual meeting of the cognitive science society (CogSci 2013)*, pp. 1438–1443 (2013)