原著論文

人間は『人工知能』と『協力』できるか：クラウドソーシングを用いた仮想的AIエージェント実験による検討

Can Humans cooperate with Artificial Intelligence?
An Examination of Virtual AI Agent Experiments Using Crowdsourcing

キーワード：
公共財ゲーム実験，人工知能，クラウドソーシング，社会的価値志向性，親密性
keyword：
Public Goods Experiment, Artificial Intelligent, Crowdsourcing, Social Value Orientation, Intimacy

明治大学　後藤　晶
Meiji University　Akira GOTO

要　約
現代社会において，人工知能（以下，AI）の開発が着々と進められており，我々の日常生活の隅々までに影響を与える時代が来ることは間違いない。そのような時代において人間がAIとのコラボレーションできるかどうかはもう一つの課題であろう。本論文においては，クラウドソーシングによる仮想的なAIエージェントを相手プレイヤーとして想定した公共財ゲーム実験を用いてAIと人間の協力行動の可能性について分析した。
その結果，公共財ゲームについては，対人信頼および対人社会的価値志向性は影響を及ぼさないが，対AI社会的価値志向性は影響を及ぼすこと，協力的なAIエージェントが多い方が公共財ゲームで貢献することが明らかになった。また，人間よりもAIに対する一般的信頼が高い一方で，向社会性と親密性についてはAIに比べて，人間に対して高く，AIエージェントに対する親密性は，一緒に作業を行っ

原稿受付：2022年８月31日
掲載決定：2023年５月22日

た後に高まることが明らかとなった。

これらの結果は，人間に対して協力的な人がAIに対して協力的であるとは限らないものの，社会的選好の知見がAIとの関係にも応用し得ること，これからの社会においてさまざまな形でAIとの協力の機会が増えることで，個人のAIに対する態度が改善し得ることを示唆している。

Abstract

Artificial intelligence (AI) is steadily developing in today's society. Someday, AI will influence every aspect of our lives, and human–AI collaboration will be a significant concern. This paper aims to explore such cooperation via public goods game experiments with virtual AI agents and participants recruited through a crowdsourcing service.

The findings revealed that trust and social value orientation toward humans did not impact the public goods game contributions. However, social value orientation toward AI did have an effect, and humans cooperated more with cooperative AI agents. Although general trust toward AI agents was higher than trust toward humans, prosociality and intimacy were still stronger toward humans. Interestingly, intimacy toward AI agents increased after engaging in public goods games with them.

These results suggest that while individuals exhibiting cooperative behavior toward humans may not necessarily display the same behavior toward AI, social preferences still can apply to human-AI relationship studies. Furthermore, enhancing human-AI cooperation opportunities is thought to improve individuals' attitudes toward AI.

1　問題

昨今では，世界中で国家をあげて，産官学民をあげて人工知能（Artificial Intelligence，以下，AI）の開発が著しく進んでいる。日本政府ではサイバー空間とフィジカル空間を高度に融合させたシステムにより，経済発展と社会的課題の解決を両立する，人間中心の社会であるSociety5.0の実現を目指している（内閣府，2016）。具体的には，IoT（Internet of Things，モノのインターネット）とAIによって必要な情報をいつでも獲得できるようになったり，ロボットや自動走行車により少子高齢化や地方の過疎，貧富の格差といった社会問題の解決が図れると考えられている。

一方で，産業界においても強力にAIの開発が進められている。いわゆるGAFA+MやFAANGと言われるようなビッグテック企業を支える技術の一つがAIであるといっても過言ではないし，我々の日常生活のさまざまなところに浸透しつつある。例えば，iPhoneに搭載されているSiriやAmazon Echoに搭載されているAlexaといった音声アシスタント機能，クレジットカードの不正使用検知，非接触型の体温検知装置などはその例であろう。それ以外にも，売上予想や医療診断，自動運転などを支える基礎技術として非常に大きな役割を果たしている。

これらのAIは人間を圧倒する情報処理能力を有しつつある。例えば，Google DeepMind（現：DeepMind Technologies）により開発されたAlpha Goは2017年５月に人類最強の棋士と称される柯潔との三番勝負で全勝するなど（WIRED, 2017）はその例であろう。このような現状を鑑みると，これからの我々の日常生活において，AIは切っても切りきれない，必要不可欠なものになりつつあることは間違いない。

ここで，改めて検討する必要があるのは，人間によるAIの受容，すなわち人間とAIの関係性であろう。これからの社会には単純に音声アシスタント機能などのAIを「使う」のであったり，Alpha GoのようなAIと「競争」したりするだけでは済まされない時代が来る。特に，今後重要となるのは人間とAIとの「協働」や「協力」である。

OECDはAIによるさまざまなメリットを得るためには「信頼できるAI（trustworthy AI，信頼されるAI）」を重要視している（OECD, 2021）。高度なAIはその技術上，ブラックボックスが存在し，説明できない要因も存在するために，他のICTに比べると信頼性が不十分である点もある。

しかし，これからの社会においてはAIと人間は共に歩んで行くのであろう。文科省における令和２年度の戦略目標及び研究開発目標において，「信頼されるAI」として「AI技術と人間の親和性が高まり，AI応用システムが人間に寄り添い，意図や文脈を理解して人々の生活・活動を適切にサポートしてくれる社会」が見据えるべき将来の社会像の一つとして掲げられている（文部科学省，2020）。そのような社会を実現するためには，人間が不確実性の存在するAIと協働・協力できるのか，どのような態度を抱くのか，といったことは課題である。

AIに対する人間の態度についてはいくつかの調査が行われている。Longoni, et. al（2019）ではAIに対して抵抗感を持っていることが指摘されている。他にも，アメリカにおいては社会におけるAIの発展が与える影響について，二分されている一方で，日本では半数以上が良いものであると捉えているなど，AIに対する意識は文化差・個人差が存在するという（Funk et al, 2020）。

実際に，総務省（2016）によれば，国内のAIに対する態度に着目すると，仕事のパートナーとしてのAIについて，上司となることに対する抵抗を持つ人は多い一方で，同僚や部下となることに対する抵抗感は多くないという。少なくとも，同等な立場で協力して作業を行うことに対してはそこまで心理的なハードルは高くないと考えられる。

消費者庁（2020）によれば，多くの消費者が「暮

らしを豊かにする」であったり，「生活に良い影響を与える」という認識である一方で，「親しみがもてる」という認識を持っている人は少ないという。

他にもMS&ADインターリスク総研株式会社の調査によれば（MS&ADインターリスク総研株式会社，2021），AIに対する期待感は高まりつつあるものの，「AIに支配される」「雇用の減少」「AIの示す情報が人為的に操作される」「想定外のAIの判断で損害が発生する」といった観点に不安に感じている人が多く，「不安に感じることはない」という回答も少なくAIに対する信頼性が低い現状がある。以上の状況を踏まえると，人間とAIの関係性については時代を経るについて期待感は高まり，社会としてのニーズは高まりつつあるものの，未だに一定の不安は存在しているようである。これらの不安はどのように改善されるのだろうか。

AIが社会において重要な役割を果たすためには，AIに対する信頼のみならず，人間と協力・協調するAI（Cooperative AI）という観点も必要不可欠である。Dafoe et. al,（2021）では，AIが成功するためには，社会科学・行動科学・自然科学といった広範な観点からAIと人間の協力に関わる研究が重要であることを指摘している。

AIと人間の協力に着目すれば，Crandallらは囚人のジレンマを用いた実験によって，人間とAIエージェント(1)同士の協力が，人間同士の協力より少ないこと，人間とAIエージェントの間に選択式のチープトークを可能にすると人間とAIエージェントの協力行動が人間同士の協力と同程度に観察されたことが報告されている（Crandall, et. al, 2018）。一方，協力行動に限らず，対人間と対コンピュータエージェントのゲーム実験での行動の差異を指摘する研究もあり，議論が割れているのが現状である（Blount, 1995；Eckel & Wilson, 2006；奥山ら，2022, Paeng et. al, 2016）。

本研究では，協力行動を分析する枠組みの一つである公共財ゲームを用いて，AIと人間の間で協力行動が観察されるか，特にどのような行動を行うAIに対して協力行動を行うのかを検討する。公共財ゲームとは，個人の利益を最大化する状況と社会の利益を最大化する状況が一致しない社会的ジレンマ状況の一つであり，自発的貢献行動，ないしは協力行動の指標として幅広く用いられているものである（Andreoni, 1995；Zelmer, 2003）。おおよそ，初期保有として与えられた金額のうち，公共財のためにいくら貢献するのか決定することが求められ，貢献額の多寡が協力行動の指標として機能する。

実験経済学や行動経済学の枠組みの中で，人間以外のエージェントとの関係について検討されたものは必ずしも多くないが，コンピュータ等の人間以外のプレイヤーによる公共財ゲーム実験研究においては，対人条件に比べて，対コンピュータ条件で貢献額が少ないことを指摘する研究がある一方で（Houser & Kurzban, 2002），対人条件と対コンピュータ条件における貢献額が相関することが指摘されるなど（Ferraro, & Vossler, 2010；Burton-Chellew, et. al, 2016），情報化社会の進展状況も影響すると考えられ，決定的な定説は存在していない。

ここでAIと人間の関係を考えるための一つのヒントとなり得るのは対人関係ないしは行動経済学における社会的選好（Social Preference）に関する研究であろう。人間対人間の関係と同様の関係性がAIと人間の間に見られるかどうかは，人間がAIと協力できるかどうか検討するための材料となりえる。実際に，機械に対する信頼研究では，人間や組織における研究手法を援用されており（笠木，2018），探索的に検討するために対人関係・社会的選好に関する手法の導入はあながち的外れではないと考えられる。本研究では，社会的選好の指標として社会的価値志向性，信頼と対他的親密性に着目する。これらの概念は対人研究の文脈で取り上げられてきたが，一緒に協力行動を行うエージェントとしてのAIに対して，人

間と同様に向社会性や信頼，親密感を抱く可能性は十分にある。

社会的価値志向性（Social Value Orientation, 以下SVO）とは，自身と他者の利得配分に対する選好を示すものである。さまざまなゲーム的な状況や現実場面での向社会的行動を説明することが指摘されており（森，2015），行動経済学における社会的選好として研究が積み重ねられているものと近接した概念である。例えば，向社会的な人は向自己的な人に比べて社会的ジレンマや公共財ゲームにおいてより協力的な傾向を示すことが指摘されている（De Cremer, and Van Lange, 2001；De Cremer and Van Dijk, E, 2002）。本研究においてはSVOスライダー法を用いて（Murphy, et. al, 2011），SVOの影響についても検討する。

信頼については，システムの動作に対する期待といったいわゆる信頼性工学的な観点ではなく，人間がAIに対してどのように認識しているのかという主観的な側面に着目する。一般的信頼とは，特定の相手を対象とせずに一般的な他者に対する信頼を示すものであり，相手に対する情報がない場合の相手の信頼性に対するデフォルトの推定値であると言える（山岸，1998）。人間とAIが協力して働く場面には信頼が影響する可能性がある。また，AIに対して示す信頼と個人に対して示す信頼についても差異が存在する可能性がある。

ただし，笠木（2018）によれば機械と人間の信頼の違いとして，（１）信頼の変化に相違があること，（２）意図をもたない機械に対する信頼は，人間に対する信頼とは異なるメカニズムにより形成され得ること，（３）人間に対する信頼は相互的に形成されるが，機械に対する信頼は一方的にならざるを得ないことが指摘されている。本研究においては，共に作業を行うエージェントとしてのAIに対する信頼を評価するため，横井・中谷内（2018）を考慮に入れ，一般的信頼尺度を参考に（Yamagishi, et. al, 2013），信頼の影響についても検討する。

対他的親密性とは他者に対する間柄の近さを示したものであり，親近感や関係性の維持を反映したものである。また，共感的関心反応と正の相関があるという（Cialdini, 1997）。本研究では，この対他的親密性を検討するためにIOS（Inclusion of Other in the Self，心理的重なり）尺度を用いた（Aron et. al, 1992）。この尺度では自身と他者が２つの円で図示されている。円が完全に離れたものから重なり合っている７種類の図を用いて，２者の関係を表すものを選ぶというものである。この尺度では円の重なりが大きいほど２者間の親密性が高いことを示している。Tatsukawa et. al,（2019）では，IOSを用いてアンドロイドに対する親密性の関係を分析しており，AIエージェントを対象に応用しても差し支えないものと考えられる。

SVO，信頼，対他的親密性のいずれも関わる知見は対人関係における研究として重要である一方で，これらの観点を拡張すれば，その他のエージェントに関する研究への応用可能性も期待できる。

本研究においては，これらの対人行動に関する知見を援用して，対AIエージェントの人間行動について以下４点が論点となる。１．人間はAIと協力できるのか，２．AIと協力しやすい人間やしにくい人間がいるのであれば，どのような要因が影響するのか，３．人間はAIに対してどのように認識しているのか，さらに４．AIと協力して作業を行うことで，共に働いたAIに対する認識がどのように変化するのか。

具体的には，協力に関わる主たる検討課題１・２と，それ以外の従たる検討課題３・４をあげる。

検討課題１：人間とAIの協力については，協力行動を分析する枠組みである公共財ゲームを用いて分析する。公共財ゲームにおけるエージェントとして，基本行動パターンとしてランダム，非協力的，中立的，協力的なAIエージェントを想定する。また，現実には，AIエージェントとの関係はさまざまな状況が想定できるが，本研究で

は，相手に対する情報がない場合に個人が一般的にイメージするAIと協力ができるのかどうか検討することを目的として，特定の文脈に依存しないコンテキストフリーな状況で実験を行う。

検討課題２：AIに対して協力的な人間の特徴については，先述のAIに対する認識に関する調査と公共財ゲームをあわせて分析することで明らかにする。

検討課題３：人間のAIに対する認識については，信頼，対他的親密性，SVOをそれぞれ対人条件及び対AI条件を比較することで，AIに対して人間がどのような認識をしているのかを明らかにする。

検討課題４：AIとの共に働く作業の影響についてはIOS尺度に着目して，公共財ゲームの事前及び事後における変化を分析する。

これらの観点を踏まえて，本研究の目的はAIと「協力」しやすい人の特徴について解明することにある。本研究では公共財ゲームを用いて，AIの行動傾向が人間の意思決定に与える影響の基礎研究となすために，「利己的なAI」「中立的なAI」「協力的なAI」「ランダムなAI」の４種類のAIエージェントを想定して，これらの仮想エージェントと人間の協力行動について検討する。

本研究のアプローチは行動経済学や社会心理学の知見を援用するアプローチであり，AIと社会や人間の関係性について分析する上で有用であると考えられる。

２　方法

2.1　実験参加者

本研究では「Yahoo!クラウドソーシング（http://crowdsourcing.yahoo.co.jp/）」を用いた。実験は2020年10月30日17時00分から23時25分にかけて実施した。実験参加者は1,138名（年齢M＝44.87, SD＝10.60，性別，年齢回答を拒否した方を除く），内訳は男性が742名（年齢M＝46.59，SD＝9.92），女性が396名（年齢M＝41.64, SD＝11.07）であった。

Yahoo!クラウドソーシングにおけるデータの代表性について簡単に述べる。本研究における実験参加者はあくまでもYahoo! Japan IDを登録している人に限られている。したがって，必ずしも十分な代表性を担保できていない可能性は存在する。しかしながら，ラボ実験のように少人数ではなく1,000人規模の実験として実施したこと，さらに実験参加者をランダムに割り当てるランダム化比較実験として実施したことにより，一定程度の妥当性を有していると考えられる。

実験参加者は固定報酬および実験結果に応じた成果報酬（最小60ポイント，最大110ポイント）をPayPayボーナスライト（現PayPayポイント）によって受け取っている。

2.2　手続き

手続きは以下の通りである。はじめに一般的な他者に対する親近感を円の重なり具合によって評価するIOS尺度について尋ねた上で（Aron et. al, 1992），同尺度により一般的なAIに対する親近感を評価した。

続いて，公共財ゲーム実験の説明を行った。なお，その際に今回の実験においては自分以外のプレイヤーがAIエージェントであると教示している。公共財ゲームの理解度確認問題の後に公共財ゲームを実施した。公共財ゲームについては，初期保有額を100ポイントとして，３人プレイヤー，毎期相手が変わらないパートナー条件を想定して実施した。この時，プレイヤーiの利得をπ_i，支払額をC_i，プレイヤー全員の支払額の合計をΣC_i，とすると利得関数は，$\pi_i=100-C_i+2/3\Sigma C_i$として表すことができる。

実験参加者には３人プレイヤーのうち，自身以外のプレイヤーがAIエージェントとして教示しているために，実験参加者はAIエージェントとプレイしていると認識していると考えられる。しかし，実際にはAIエージェントの貢献額は，一

定範囲内でランダムに貢献額を決定するように操作している。０–33ポイントをランダムに貢献する利己的仮想エージェント，34–66ポイントをランダムに貢献する中立的仮想エージェント，67–100ポイントをランダムに貢献する協力的仮想エージェント，さらに０–100ポイントをランダムに貢献するランダム仮想エージェントを設定し，自身以外の２プレイヤーの４仮想エージェントの組み合わせで４×４の16条件について実験を行った。

公共財ゲーム実験を実施した後に，同実験で参加した仮想エージェントそれぞれに対する親近感をIOS尺度により尋ねた（Aron et. al, 1992)。その他の調査項目として一般的信頼尺度(Yamagishi, et. al, 2013）と同尺度をもとに作成した対AI一般的信頼尺度，社会的価値志向性を明らかにするSVOスライダー（Murphy et. al, 2011）と同実験をもとに作成した対AI SVOスライダー，３問の設問からなる認知能力を評価する認知反射テスト（Fredrick, 2005）および性別・年齢・居住地域等の社会経済的要因に関する項目を尋ねた。これらの実験・調査システムはoTreeを用いて開発した（Chen, et. al, 2016；後藤, 2021)。

2.3　分析手法

はじめに，公共財ゲームにおける貢献行動の分析を行う。今回の公共財ゲーム実験においては，一人の実験参加者が複数期参加していることから，いわゆる重回帰分析等の一般線形モデルによる分析の実施は適切ではない。また，今回は公共財ゲームにおける分配の割合に注目することを踏まえて，一般化線形混合モデルのロジスティック回帰分析モデルにより分析を行う。また，実験条件の説明変数には各条件における仮想エージェントの個数を設定する。具体的には利己的仮想エージェント，中立的仮想エージェント，協力的仮想エージェントの個数を説明変数として投入する。さらに，社会経済的要因やN–1期目における貢献額等を踏まえて，その協力行動の差異を明らかにする。

続いて，調査項目に関する比較として，一般的信頼尺度，SVOスライダー，IOS尺度について対人間条件と対AI条件の２群の平均値の差の分析を対応のあるt検定によって行い，人間に対する態度とAIに対する態度を比較する。なお，一般的信頼尺度においては，笠木（2018）の指摘する信頼の一方向性を考慮して分析を行う[(2)]。

最後に，協力行動を反映する公共財ゲーム実験の前と後で，AIに対する態度がどのように変化するか検討するために，事前の一般的なAIに対するIOS尺度と，２体の公共財ゲーム実験に参加した仮想エージェントに対するIOS尺度の値を，一般線形混合モデルにより比較する。

3　結果

3.1　記述統計量

本研究で得たデータの記述統計量は表１に示した通りである。各条件について，最小で57名，最大で78名のデータとなっているが，各項目に大きな偏りはなく，分析において大きな問題はないと考えられる。

今回の実験においては，おおよそ40代の実験参加者が多く，続いて50代，30代の順番であった。また，個人収入については１–２百万円台の方が多く，続いて２–４百万円台，４–６百万円台の方の順であった。居住地域については，関東，近畿，中部の順番であり，既婚者と未婚者はほぼ同数，子どものいない人が６割前後を占めている。

3.2　公共財ゲームに関する分析

続いて，本研究の主たる結果である公共財ゲームの結果に関する分析を示す。図１には各条件における平均貢献額のプロットを示し，分析結果を表２に示す。これにより，検討課題１および２について明らかにする。

表1　実験条件ごとの記述統計量

エージェント1	利己的				中立的				向社会的				ランダム				全体
エージェント2	利己的	中立的	向社会的	ランダム	利己的	中立的	向社会的	ランダム	利己的	中立的	向社会的	ランダム	利己的	中立的	向社会的	ランダム	
	(Obs.=780)	(Obs.=720)	(Obs.=570)	(Obs.=760)	(Obs.=760)	(Obs.=750)	(Obs.=620)	(Obs.=660)	(Obs.=670)	(Obs.=760)	(Obs.=760)	(Obs.=640)	(Obs.=660)	(Obs.=750)	(Obs.=740)	(Obs.=780)	(Obs.=11380)
賞金額（10種）																	
Mean (SD)	41.8 (33.3)	39.7 (31.7)	53.2 (31.9)	45.2 (33.7)	37.5 (30.9)	45.4 (33.3)	54.2 (30.7)	47.5 (30.2)	40.9 (31.3)	51.0 (31.6)	49.4 (32.8)	45.6 (35.2)	42.4 (33.1)	50.1 (31.1)	50.7 (32.3)	42.8 (31.4)	45.9 (32.5)
Median [Min, Max]	40.0 [0, 100]	39.0 [0, 100]	50.0 [0, 100]	50.0 [0, 100]	33.0 [0, 100]	44.0 [0, 100]	51.0 [0, 100]	50.0 [0, 100]	40.0 [0, 100]	50.0 [0, 100]	50.0 [0, 100]	50.0 [0, 100]	39.0 [0, 100]	50.0 [0, 100]	50.0 [0, 100]	50.0 [0, 100]	50.0 [0, 100]
	(N=78)	(N=72)	(N=57)	(N=76)	(N=76)	(N=75)	(N=62)	(N=66)	(N=67)	(N=76)	(N=76)	(N=64)	(N=66)	(N=75)	(N=74)	(N=78)	(N=1138)
利己的仮想エージェント個数																	
Mean (SD)	2.00 (0)	1.00 (0)	1.00 (0)	1.00 (0)	1.00 (0)	0 (0)	0 (0)	0 (0)	1.00 (0)	0 (0)	0 (0)	0 (0)	1.00 (0)	0 (0)	0 (0)	0 (0)	0.501 (0.622)
Median [Min, Max]	2.00 [2.00, 2.00]	1.00 [1.00, 1.00]	1.00 [1.00, 1.00]	1.00 [1.00, 1.00]	1.00 [1.00, 1.00]	0 [0, 0]	0 [0, 0]	0 [0, 0]	1.00 [1.00, 1.00]	0 [0, 0]	0 [0, 0]	0 [0, 0]	1.00 [1.00, 1.00]	0 [0, 0]	0 [0, 0]	0 [0, 0]	0 [0, 2.00]
中立的仮想エージェント個数																	
Mean (SD)	0 (0)	1.00 (0)	0 (0)	0 (0)	1.00 (0)	2.00 (0)	1.00 (0)	1.00 (0)	0 (0)	1.00 (0)	0 (0)	0 (0)	0 (0)	1.00 (0)	0 (0)	0 (0)	0.507 (0.618)
Median [Min, Max]	0 [0, 0]	1.00 [1.00, 1.00]	0 [0, 0]	0 [0, 0]	1.00 [1.00, 1.00]	2.00 [2.00, 2.00]	1.00 [1.00, 1.00]	1.00 [1.00, 1.00]	0 [0, 0]	1.00 [1.00, 1.00]	0 [0, 0]	0 [0, 0]	0 [0, 0]	1.00 [1.00, 1.00]	0 [0, 0]	0 [0, 0]	0 [0, 2.00]
協力的仮想エージェント個数																	
Mean (SD)	0 (0)	0 (0)	1.00 (0)	0 (0)	0 (0)	0 (0)	1.00 (0)	0 (0)	1.00 (0)	1.00 (0)	2.00 (0)	1.00 (0)	0 (0)	0 (0)	1.00 (0)	0 (0)	0.485 (0.619)
Median [Min, Max]	0 [0, 0]	0 [0, 0]	1.00 [1.00, 1.00]	0 [0, 0]	0 [0, 0]	0 [0, 0]	1.00 [1.00, 1.00]	0 [0, 0]	1.00 [1.00, 1.00]	1.00 [1.00, 1.00]	2.00 [2.00, 2.00]	1.00 [1.00, 1.00]	0 [0, 0]	0 [0, 0]	1.00 [1.00, 1.00]	0 [0, 0]	0 [0, 2.00]
確認テスト正答数																	
Mean (SD)	4.72 (2.10)	5.35 (2.05)	5.07 (1.88)	4.97 (1.90)	5.30 (1.83)	5.37 (1.86)	4.98 (1.91)	5.20 (1.98)	5.49 (1.52)	4.96 (2.13)	5.13 (2.02)	5.03 (2.23)	5.24 (1.84)	5.44 (1.83)	5.58 (1.90)	5.28 (1.90)	5.20 (1.94)
Median [Min, Max]	5.00 [1.00, 8.00]	6.00 [1.00, 8.00]	5.00 [1.00, 8.00]	5.00 [1.00, 8.00]	6.00 [1.00, 8.00]	5.00 [1.00, 8.00]	5.00 [1.00, 8.00]	6.00 [1.00, 8.00]	6.00 [2.00, 8.00]	5.00 [1.00, 8.00]	5.00 [0, 8.00]	5.50 [0, 8.00]	5.00 [1.00, 8.00]	6.00 [1.00, 8.00]	6.00 [1.00, 8.00]	5.00 [0, 8.00]	5.00 [0, 8.00]
対AIIOS																	
Mean (SD)	2.38 (1.45)	2.31 (1.35)	2.56 (1.70)	2.57 (1.52)	2.25 (1.40)	2.44 (1.59)	2.65 (1.67)	2.30 (1.34)	2.36 (1.44)	2.49 (1.54)	2.32 (1.46)	2.70 (1.49)	2.65 (1.54)	2.68 (1.59)	2.61 (1.65)	2.78 (1.61)	2.50 (1.52)
Median [Min, Max]	2.00 [1.00, 5.00]	2.00 [1.00, 6.00]	2.00 [1.00, 7.00]	2.00 [1.00, 6.00]	2.00 [1.00, 5.00]	2.00 [1.00, 7.00]	2.00 [1.00, 7.00]	2.00 [1.00, 6.00]	2.00 [1.00, 6.00]	2.00 [1.00, 6.00]	2.00 [1.00, 7.00]	2.50 [1.00, 6.00]	3.00 [1.00, 6.00]	2.00 [1.00, 7.00]	2.50 [1.00, 6.00]	2.00 [1.00, 7.00]	2.00 [1.00, 7.00]
対人IOS																	
Mean (SD)	3.04 (1.70)	2.93 (1.61)	2.79 (1.59)	3.41 (1.86)	3.20 (1.60)	3.19 (1.84)	3.35 (1.81)	3.05 (1.59)	3.15 (1.69)	3.22 (1.74)	3.42 (1.78)	3.42 (1.74)	3.32 (1.62)	3.37 (1.50)	3.38 (1.88)	3.28 (1.62)	3.22 (1.70)
Median [Min, Max]	2.50 [1.00, 7.00]	3.00 [1.00, 6.00]	3.00 [1.00, 6.00]	3.50 [1.00, 7.00]	3.00 [1.00, 7.00]	3.00 [1.00, 7.00]	4.00 [1.00, 7.00]	3.00 [1.00, 6.00]	3.00 [1.00, 7.00]	3.00 [1.00, 7.00]	4.00 [1.00, 7.00]	3.00 [1.00, 7.00]	3.50 [1.00, 6.00]	4.00 [1.00, 7.00]	3.00 [1.00, 7.00]	3.00 [1.00, 7.00]	3.00 [1.00, 7.00]
対仮想エージェントAIOS																	
Mean (SD)	2.54 (1.69)	2.57 (1.60)	3.32 (1.85)	2.80 (1.78)	3.22 (1.59)	2.93 (1.63)	3.32 (2.03)	3.32 (1.67)	3.10 (1.75)	3.38 (1.81)	3.03 (1.62)	3.16 (1.83)	3.38 (1.39)	3.03 (1.68)	3.20 (1.91)	3.06 (1.60)	3.07 (1.72)
Median [Min, Max]	2.00 [1.00, 7.00]	2.00 [1.00, 6.00]	3.00 [1.00, 7.00]	2.00 [1.00, 7.00]	3.00 [1.00, 6.00]	3.00 [1.00, 6.00]	3.00 [1.00, 7.00]	3.00 [1.00, 7.00]	3.00 [1.00, 7.00]	3.00 [1.00, 7.00]	3.00 [1.00, 7.00]	3.00 [1.00, 7.00]	3.50 [1.00, 6.00]	3.00 [1.00, 7.00]	3.00 [1.00, 7.00]	3.00 [1.00, 7.00]	3.00 [1.00, 7.00]
対仮想エージェントBIOS																	
Mean (SD)	2.45 (1.58)	3.11 (1.45)	3.28 (1.74)	2.86 (1.67)	2.83 (1.72)	2.92 (1.69)	3.47 (1.86)	3.02 (1.72)	2.93 (1.86)	3.32 (1.69)	2.99 (1.64)	3.20 (1.84)	3.02 (1.41)	3.19 (1.55)	3.26 (1.73)	2.91 (1.35)	3.03 (1.66)
Median [Min, Max]	2.00 [1.00, 7.00]	3.00 [1.00, 6.00]	3.00 [1.00, 7.00]	3.00 [1.00, 7.00]	2.50 [1.00, 7.00]	3.00 [1.00, 6.00]	3.00 [1.00, 7.00]	3.00 [1.00, 7.00]	3.00 [1.00, 7.00]	3.00 [1.00, 7.00]	3.00 [1.00, 7.00]	3.00 [1.00, 7.00]	3.00 [1.00, 6.00]	3.00 [1.00, 7.00]	3.00 [1.00, 7.00]	3.00 [1.00, 6.00]	3.00 [1.00, 7.00]
対AI一般的価値																	
Mean (SD)	15.2 (4.67)	15.3 (3.92)	15.5 (4.60)	14.9 (4.31)	15.3 (4.10)	17.0 (3.20)	17.0 (3.53)	17.2 (3.76)	15.7 (4.45)	16.5 (3.57)	16.9 (4.97)	16.9 (3.45)	16.1 (3.38)	15.6 (3.96)	16.5 (3.41)	16.5 (3.52)	16.1 (4.03)
Median [Min, Max]	15.5 [4.00, 28.0]	16.0 [5.00, 24.0]	16.0 [4.00, 26.0]	15.5 [4.00, 28.0]	16.0 [4.00, 28.0]	16.0 [8.00, 28.0]	17.0 [4.00, 28.0]	17.0 [8.00, 28.0]	16.0 [4.00, 24.0]	16.0 [8.00, 24.0]	17.0 [4.00, 28.0]	17.0 [4.00, 24.0]	16.0 [4.00, 23.0]	16.0 [4.00, 26.0]	17.0 [7.00, 23.0]	17.0 [7.00, 25.0]	16.0 [4.00, 28.0]
対人価値																	
Mean (SD)	15.0 (4.08)	14.7 (4.52)	15.1 (4.96)	14.9 (4.84)	14.9 (4.49)	15.9 (4.77)	15.5 (4.13)	16.2 (3.90)	14.0 (4.51)	14.5 (4.81)	15.1 (5.24)	16.3 (4.00)	15.4 (4.23)	14.7 (4.96)	15.2 (4.18)	15.0 (4.56)	15.1 (4.57)
Median [Min, Max]	15.0 [4.00, 28.0]	15.5 [4.00, 24.0]	16.0 [4.00, 25.0]	16.0 [4.00, 24.0]	15.0 [4.00, 25.0]	17.0 [4.00, 28.0]	15.5 [7.00, 28.0]	16.0 [7.00, 24.0]	14.0 [4.00, 23.0]	15.0 [4.00, 24.0]	16.0 [4.00, 26.0]	16.0 [5.00, 24.0]	15.5 [4.00, 24.0]	15.0 [4.00, 25.0]	15.0 [7.00, 24.0]	16.0 [4.00, 24.0]	16.0 [4.00, 28.0]
対AISVO																	
Mean (SD)	20.5 (12.0)	20.9 (12.1)	20.5 (12.2)	23.5 (11.7)	21.7 (12.5)	22.6 (9.08)	21.4 (9.67)	22.3 (10.1)	24.9 (13.2)	23.5 (10.9)	21.7 (12.4)	22.4 (13.4)	23.6 (11.0)	24.6 (12.5)	22.1 (9.80)	22.3 (11.4)	22.4 (11.6)
Median [Min, Max]	22.6 [-15.1, 45.0]	22.6 [-9.94, 51.0]	22.6 [-16.3, 42.3]	22.9 [-7.82, 45.0]	22.6 [-7.82, 45.5]	22.6 [5.63, 39.2]	22.6 [-16.3, 36.8]	22.6 [-7.82, 40.8]	23.9 [-7.82, 61.4]	22.6 [-7.82, 45.0]	22.6 [-7.82, 45.0]	22.6 [-9.24, 61.4]	22.6 [-6.93, 53.4]	22.6 [-7.88, 58.4]	22.6 [-7.82, 41.3]	23.0 [-7.82, 43.1]	22.6 [-16.3, 61.4]
対人SVO																	
Mean (SD)	21.8 (11.7)	23.4 (12.1)	23.8 (11.9)	24.6 (12.1)	26.0 (11.5)	23.8 (10.5)	26.2 (10.9)	25.0 (10.0)	25.1 (13.3)	27.1 (9.61)	23.2 (12.6)	25.0 (13.4)	23.1 (10.6)	25.1 (11.8)	24.6 (9.80)	24.2 (12.1)	24.5 (11.5)
Median [Min, Max]	22.6 [-7.82, 46.4]	22.7 [-11.4, 60.9]	24.4 [-15.2, 38.4]	22.8 [-9.27, 45.0]	26.9 [-7.82, 45.8]	22.7 [-7.82, 37.4]	24.1 [0, 45.4]	22.7 [1.77, 45.0]	24.7 [-16.3, 61.4]	25.8 [-0.770, 45.0]	22.6 [-7.82, 45.0]	25.0 [-7.48, 61.4]	22.6 [-7.82, 45.2]	24.2 [-7.40, 59.3]	22.6 [-7.82, 39.1]	23.9 [-7.82, 40.3]	22.8 [-16.3, 61.4]
認知反射テスト																	
Mean (SD)	1.38 (1.18)	1.58 (1.17)	1.35 (1.16)	1.70 (1.10)	1.61 (1.19)	1.79 (1.11)	1.42 (1.21)	1.48 (1.22)	1.63 (1.19)	1.58 (1.12)	1.57 (1.18)	1.52 (1.07)	1.45 (1.07)	1.56 (1.09)	1.61 (1.18)	1.40 (1.12)	1.54 (1.15)
Median [Min, Max]	1.00 [0, 3.00]	2.00 [0, 3.00]	2.00 [0, 3.00]	2.00 [0, 3.00]	2.00 [0, 3.00]	2.00 [0, 3.00]	1.50 [0, 3.00]	1.00 [0, 3.00]	2.00 [0, 3.00]	2.00 [0, 3.00]	2.00 [0, 3.00]	2.00 [0, 3.00]	1.50 [0, 3.00]	2.00 [0, 3.00]	2.00 [0, 3.00]	1.00 [0, 3.00]	2.00 [0, 3.00]
性別																	
男性	54 (69.2%)	37 (51.4%)	35 (61.4%)	52 (68.4%)	47 (61.8%)	53 (70.7%)	39 (62.9%)	39 (59.1%)	47 (70.1%)	56 (73.7%)	47 (61.8%)	45 (70.3%)	42 (63.6%)	55 (73.3%)	45 (60.8%)	49 (62.8%)	742 (65.2%)
女性	24 (30.8%)	35 (48.6%)	22 (38.6%)	24 (31.6%)	29 (38.2%)	22 (29.3%)	23 (37.1%)	27 (40.9%)	20 (29.9%)	20 (26.3%)	29 (38.2%)	19 (29.7%)	24 (36.4%)	20 (26.7%)	29 (39.2%)	29 (37.2%)	396 (34.8%)
世代																	
10代	0 (0%)	1 (1.4%)	1 (1.8%)	1 (1.3%)	1 (1.3%)	0 (0%)	0 (0%)	0 (0%)	1 (1.5%)	0 (0%)	3 (3.9%)	1 (1.6%)	0 (0%)	0 (0%)	1 (1.4%)	0 (0%)	10 (0.9%)
20代	5 (6.4%)	5 (6.9%)	5 (8.8%)	3 (3.9%)	3 (3.9%)	7 (9.3%)	0 (0%)	2 (3.0%)	3 (4.5%)	7 (9.2%)	3 (3.9%)	3 (4.7%)	2 (3.0%)	10 (13.3%)	7 (9.5%)	9 (11.5%)	74 (6.5%)
30代	15 (19.2%)	13 (18.1%)	15 (26.3%)	14 (18.4%)	23 (30.3%)	15 (20.0%)	13 (21.0%)	10 (15.2%)	21 (31.3%)	15 (19.7%)	18 (23.7%)	17 (26.6%)	16 (24.2%)	19 (25.3%)	18 (24.3%)	14 (17.9%)	256(22.5%)
40代	34 (43.6%)	30 (41.7%)	24 (42.1%)	22 (28.9%)	22 (28.9%)	26 (34.7%)	30 (48.4%)	30 (45.5%)	27 (40.3%)	32 (42.1%)	25 (32.9%)	22 (34.4%)	30 (45.5%)	24 (32.0%)	31 (41.9%)	30 (38.5%)	439 (38.6%)
50代	15 (19.2%)	13 (18.1%)	10 (17.5%)	25 (32.9%)	21 (27.6%)	26 (34.7%)	13 (21.0%)	20 (30.3%)	12 (17.9%)	19 (25.0%)	21 (27.6%)	14 (21.9%)	11 (16.7%)	15 (20.0%)	12 (16.2%)	20 (25.6%)	267 (23.5%)
60代	7 (9.0%)	8 (11.1%)	1 (1.8%)	10 (13.2%)	6 (7.9%)	1 (1.3%)	4 (6.5%)	2 (3.0%)	2 (3.0%)	2 (2.6%)	4 (5.3%)	4 (6.3%)	6 (9.1%)	7 (9.3%)	5 (6.8%)	4 (5.1%)	73(6.4%)
70代以上	2 (2.6%)	2 (2.8%)	1 (1.8%)	1 (1.3%)	0 (0%)	0 (0%)	2 (3.2%)	2 (3.0%)	1 (1.5%)	1 (1.3%)	2 (2.6%)	3 (4.7%)	1 (1.5%)	0 (0%)	0 (0%)	1 (1.3%)	19 (1.7%)
個人収入																	
0円	7 (9.0%)	12 (16.7%)	2 (3.5%)	5 (6.6%)	7 (9.2%)	11 (14.7%)	8 (12.9%)	6 (9.1%)	10 (14.9%)	7 (9.2%)	7 (9.2%)	3 (4.7%)	4 (6.1%)	5 (6.7%)	9 (12.2%)	5 (6.4%)	108 (9.5%)
1-2百万	15 (19.2%)	19 (26.4%)	18 (31.6%)	15 (19.7%)	20 (26.3%)	19 (25.3%)	15 (24.2%)	20 (30.3%)	16 (23.9%)	14 (18.4%)	21 (27.6%)	15 (23.4%)	13 (19.7%)	12 (16.0%)	19 (25.7%)	19 (24.4%)	270 (23.7%)
2-4百万	13 (16.7%)	13 (18.1%)	10 (17.5%)	10 (13.2%)	16 (21.1%)	15 (20.0%)	13 (21.0%)	8 (12.1%)	18 (26.9%)	15 (19.7%)	11 (14.5%)	16 (25.0%)	15 (22.7%)	23 (30.7%)	15 (20.3%)	13 (16.7%)	224 (19.7%)
4-6百万	13 (16.7%)	9 (12.5%)	9 (15.8%)	18 (23.7%)	4 (5.3%)	9 (12.0%)	12 (19.4%)	14 (21.2%)	7 (10.4%)	7 (9.2%)	14 (18.4%)	11 (17.2%)	14 (21.2%)	9 (12.0%)	7 (9.5%)	14 (17.9%)	171 (15.0%)
6-8百万	9 (11.5%)	5 (6.9%)	7 (12.3%)	8 (10.5%)	8 (10.5%)	10 (13.3%)	4 (6.5%)	6 (9.1%)	5 (7.5%)	9 (11.8%)	9 (11.8%)	4 (6.3%)	9 (13.6%)	8 (10.7%)	4 (5.4%)	9 (11.5%)	114 (10.0%)
8-10百万	3 (3.8%)	2 (2.8%)	0 (0%)	2 (2.6%)	4 (5.3%)	1 (1.3%)	1 (1.6%)	6 (9.1%)	2 (3.0%)	5 (6.6%)	2 (2.6%)	2 (3.1%)	1 (1.5%)	6 (8.0%)	6 (8.1%)	2 (2.6%)	45 (4.0%)
10百万以上	4 (5.1%)	2 (2.8%)	2 (3.5%)	3 (3.9%)	4 (5.3%)	2 (2.7%)	3 (4.8%)	0 (0%)	3 (4.5%)	5 (6.6%)	1 (1.3%)	3 (4.7%)	3 (4.5%)	1 (1.3%)	3 (4.1%)	3 (3.8%)	42 (3.7%)
不明or答えたくない	14 (17.9%)	10 (13.9%)	9 (15.8%)	15 (19.7%)	13 (17.1%)	8 (10.7%)	6 (9.7%)	6 (9.1%)	6 (9.0%)	14 (18.4%)	11 (14.5%)	10 (15.6%)	7 (10.6%)	11 (14.7%)	11 (14.9%)	13 (16.7%)	164 (14.4%)
居住地域																	
関東	30 (38.5%)	25 (34.7%)	21 (36.8%)	21 (27.6%)	32 (42.1%)	29 (38.7%)	25 (40.3%)	16 (24.2%)	20 (29.9%)	37 (48.7%)	33 (43.4%)	20 (31.3%)	29 (43.9%)	27 (36.0%)	27 (36.5%)	27 (34.6%)	419 (36.8%)
北海道	2 (2.6%)	1 (1.4%)	2 (3.5%)	1 (1.3%)	2 (2.6%)	3 (4.0%)	2 (3.2%)	2 (3.0%)	5 (7.5%)	1 (1.3%)	1 (1.3%)	1 (1.6%)	6 (9.1%)	5 (6.7%)	3 (4.1%)	4 (5.1%)	41 (3.6%)
東北	4 (5.1%)	7 (9.7%)	5 (8.8%)	7 (9.2%)	3 (3.9%)	4 (5.3%)	1 (1.6%)	8 (12.1%)	2 (3.0%)	8 (10.5%)	4 (5.3%)	7 (10.9%)	3 (4.5%)	6 (8.0%)	2 (2.7%)	3 (3.8%)	74 (6.5%)
中部	10 (12.8%)	15 (20.8%)	8 (14.0%)	14 (18.4%)	14 (18.4%)	13 (17.3%)	13 (21.0%)	10 (15.2%)	14 (20.9%)	13 (17.1%)	7 (9.2%)	12 (18.8%)	10 (15.2%)	22 (29.3%)	14 (18.9%)	16 (20.5%)	205 (18.0%)
近畿	23 (29.5%)	14 (19.4%)	12 (21.1%)	16 (21.1%)	13 (17.1%)	12 (16.0%)	14 (22.6%)	17 (25.8%)	13 (19.4%)	11 (14.5%)	13 (17.1%)	13 (20.3%)	7 (10.6%)	5 (6.7%)	18 (24.3%)	16 (20.5%)	217 (19.1%)
中国	3 (3.8%)	3 (4.2%)	3 (5.3%)	7 (9.2%)	3 (3.9%)	6 (8.0%)	3 (4.8%)	3 (4.5%)	9 (13.4%)	2 (2.6%)	3 (3.9%)	4 (6.3%)	8 (12.1%)	5 (6.7%)	1 (1.4%)	2 (2.6%)	65 (5.7%)
四国	2 (2.6%)	1 (1.4%)	2 (3.5%)	3 (3.9%)	3 (3.9%)	2 (2.7%)	0 (0%)	5 (7.6%)	1 (1.5%)	1 (1.3%)	3 (3.9%)	1 (1.6%)	1 (1.5%)	0 (0%)	1 (1.4%)	3 (3.8%)	29 (2.5%)
九州	4 (5.1%)	6 (8.3%)	4 (7.0%)	7 (9.2%)	6 (7.9%)	6 (8.0%)	4 (6.5%)	5 (7.6%)	3 (4.5%)	3 (3.9%)	12 (15.8%)	6 (9.4%)	2 (3.0%)	5 (6.7%)	8 (10.8%)	7 (9.0%)	88 (7.7%)
答えたくない	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
未既婚																	
未婚	31 (39.7%)	32 (44.4%)	30 (52.6%)	36 (47.4%)	37 (48.7%)	43 (57.3%)	29 (46.8%)	28 (42.4%)	39 (58.2%)	38 (50.0%)	38 (50.0%)	31 (48.4%)	29 (43.9%)	42 (56.0%)	35 (47.3%)	39 (50.0%)	557 (48.9%)
既婚	43 (55.1%)	38 (52.8%)	25 (43.9%)	39 (51.3%)	37 (48.7%)	32 (42.7%)	31 (50.0%)	37 (56.1%)	27 (40.3%)	35 (46.1%)	37 (48.7%)	32 (50.0%)	35 (53.0%)	31 (41.3%)	33 (44.6%)	36 (46.2%)	548 (48.2%)
答えたくない	4 (5.1%)	2 (2.8%)	2 (3.5%)	1 (1.3%)	2 (2.6%)	0 (0%)	2 (3.2%)	1 (1.5%)	1 (1.5%)	3 (3.9%)	1 (1.3%)	1 (1.6%)	2 (3.0%)	2 (2.7%)	6 (8.1%)	3 (3.8%)	33 (2.9%)
子の有無																	
子なし	39 (50.0%)	35 (48.6%)	37 (64.9%)	36 (47.4%)	47 (61.8%)	47 (62.7%)	32 (51.6%)	39 (59.1%)	44 (65.7%)	43 (56.6%)	47 (61.8%)	37 (57.8%)	35 (53.0%)	45 (60.0%)	48 (64.9%)	39 (50.0%)	650 (57.1%)
子あり	36 (46.2%)	36 (50.0%)	18 (31.6%)	39 (51.3%)	26 (34.2%)	27 (36.0%)	29 (46.8%)	26 (39.4%)	23 (34.3%)	29 (38.2%)	28 (36.8%)	24 (37.5%)	29 (43.9%)	27 (36.0%)	22 (29.7%)	36 (46.2%)	455 (40.0%)
答えたくない	3 (3.8%)	1 (1.4%)	2 (3.5%)	1 (1.3%)	3 (3.9%)	1 (1.3%)	1 (1.6%)	1 (1.5%)	0 (0%)	4 (5.3%)	1 (1.3%)	3 (4.7%)	2 (3.0%)	3 (4.0%)	4 (5.4%)	3 (3.8%)	33 (2.9%)

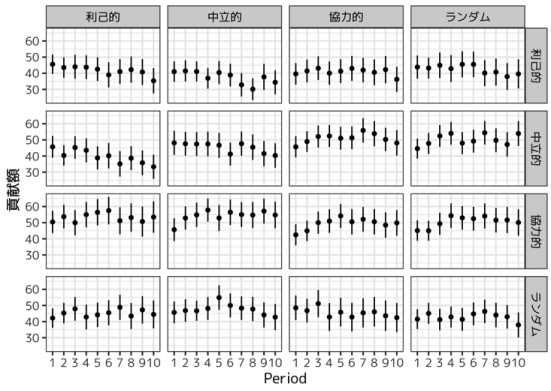

図１　各条件における貢献額の平均値

Model１は繰り返しの期数，チェック問題の正答数に加えて，対AI一般的信頼尺度，対人一般的信頼尺度，対AI SVO，対人SVOならびに認知反射テストのスコアを投入したものである。その他，社会経済的要因として性別・個人年収・居住地域・未既婚・子どもの有無を投入している。このモデルからは，期を経る毎に貢献額が減少すること，チェック問題の正答数が多く，公共財ゲームのルールを理解している人ほど貢献額が減少していること，さらに認知反射テストのスコアが高いほど貢献額が減少することが示されている。また，一般的に対人SVOは対人公共財ゲームにおいてはポジティブな影響が示されるものであるが，本研究においては，対人SVOは影響が認められず，対AISVOによるポジティブな影響が示された。

続いて，Model２はModel１から対人IOS，対人信頼，ならびに対人SVOを除いたモデルである。固定効果による分散説明率（表中固定R^2）ならびに固定効果と変量効果による分散説明率（表中固定・変量R^2）については，後者は変わらず，前者はわずかに悪化している。しかしながら，AICに着目するとModel１に比べて大きく改善していることから，Model１に比べてModel２の方が，妥当性の高いモデルでとして評価できる。また，Model１とModel２の結果を踏まえると，対人IOS，一般的信頼，SVOはいずれも影響しないことが示唆されている。以下では，Model２をベースとして分析を行う。

Model３はModel２に加えて，実験条件における各仮想エージェントの個数を説明変数に加えたものである。固定R^2は改善しているものの，固定・変量R^2については，後者はわずかに悪化している。AICは大きく改善していることからModel２に比べてModel３の方が説明力が高い。このモデルでは協力的仮想エージェントが多いことで人間が協力するという結果が示されている。

さらに，Model４およびModel５では事前の期における仮想エージェントの行動の影響を分析するため，N–1期目の２体の仮想エージェントによる貢献額の平均値を説明変数に加えて分析を行う。１期目においては０期目の貢献額が存在しないために応答変数から除いている。したがって，応答変数は２–10期におけるプレイヤーの貢献額である。Model４およびModel５のいずれもがN–1期の仮想エージェント平均貢献額とN期におけるプレイヤーの貢献額が正の相関をすることが示されている。特に，Model５ではModel３と同様に協力的な仮想エージェントが多いことで，人間がより協力的になるという結果が示されている。

なお，いずれのモデルについても説明変数のVIFは５未満であることを確認しており，多重共線性の問題は発生していない。

3.3　調査項目に関する比較

以下では，対人間条件と対AI条件について尋ねている社会的価値志向性，信頼，対他的親密性について平均値の比較を行う。これにより，検討課題３について明らかにする。

3.3.1　社会的価値志向性に関する比較

図２ではSVOスライダーの平均値について，対AI条件と対人間条件の95%信頼区間を示している。この２群において対応のあるt検定を行ったところ，$t(1137)=6.85$, $p<.001$ ($r=.20$) で人間条件の方が高いことが明らかとなった。このこ

表２　分析結果

	Model 1		Model 2		Model 3		Model 4		Model 5	
Predictors	*Odds Ratios*	*p*	*Odds Ratios*	*p*	*Odds Ratios*	*p*	*Odds Ratios*	*p*	*Odds Ratios*	*p*
(Intercept)	0.673 (0.258 – 1.759)	0.42	0.747 (0.300 – 1.863)	0.532	0.640 (0.240 – 1.708)	0.373	0.740 (0.282 – 1.945)	0.542	0.591 (0.215 – 1.625)	0.308
利己的エージェントの個数					0.976 (0.760 – 1.252)	0.847			1.050 (0.802 – 1.375)	0.721
中立的エージェントの個数					1.247 (0.974 – 1.597)	0.08			1.271 (0.973 – 1.661)	0.079
協力的エージェントの個数					1.410 ** (1.098 – 1.811)	**0.007**			1.402 * (1.070 – 1.838)	**0.014**
N-1期のエージェント平均貢献額							1.004 *** (1.003 – 1.004)	**<0.001**	1.004 *** (1.003 – 1.004)	**<0.001**
期	0.991 *** (0.989 – 0.992)	**<0.001**	0.991 *** (0.989 – 0.992)	**<0.001**	0.991 *** (0.989 – 0.992)	**<0.001**	0.981 *** (0.980 – 0.983)	**<0.001**	0.981 *** (0.980 – 0.983)	**<0.001**
確認テストの正解数	0.834 *** (0.775 – 0.898)	**<0.001**	0.835 *** (0.776 – 0.899)	**<0.001**	0.837 *** (0.778 – 0.901)	**<0.001**	0.835 *** (0.771 – 0.904)	**<0.001**	0.838 *** (0.774 – 0.907)	**<0.001**
対AIIOS尺度	0.985 (0.898 – 1.081)	0.754	0.975 (0.896 – 1.061)	0.562	0.979 (0.900 – 1.065)	0.624	0.982 (0.896 – 1.076)	0.692	0.987 (0.901 – 1.081)	0.779
対人IOS尺度	0.969 (0.888 – 1.056)	0.473								
対AI一般的信頼	1.011 (0.982 – 1.040)	0.46	1.016 (0.990 – 1.043)	0.223	1.011 (0.984 – 1.038)	0.428	1.019 (0.991 – 1.049)	0.181	1.015 (0.986 – 1.044)	0.313
対人一般的信頼	1.013 (0.987 – 1.040)	0.335								
対AISVO	1.031 *** (1.017 – 1.046)	**<0.001**	1.035 *** (1.024 – 1.047)	**<0.001**	1.035 *** (1.024 – 1.046)	**<0.001**	1.035 *** (1.022 – 1.047)	**<0.001**	1.035 *** (1.023 – 1.047)	**<0.001**
対人SVO	1.006 (0.991 – 1.020)	0.441								
認知反射テスト	0.864 * (0.761 – 0.980)	**0.023**	0.868 * (0.765 – 0.984)	**0.027**	0.863 * (0.761 – 0.978)	**0.021**	0.846 * (0.739 – 0.970)	**0.016**	0.840 * (0.733 – 0.962)	**0.012**
社会経済的要因										
性別（参照群：男性）										
女性ダミー	0.918 (0.668 – 1.260)	0.595	0.930 (0.678 – 1.275)	0.650	0.949 (0.693 – 1.299)	0.742	0.905 (0.643 – 1.273)	0.566	0.922 (0.656 – 1.295)	0.637
世代（参照群：20代）										
10代ダミー	1.030 (0.248 – 4.274)	0.968	0.980 (0.234 – 4.105)	0.978	0.869 (0.210 – 3.599)	0.846	0.852 (0.183 – 3.965)	0.838	0.762 (0.162 – 3.581)	0.73
30代ダミー	1.194 (0.670 – 2.128)	0.548	1.164 (0.657 – 2.062)	0.603	1.147 (0.647 – 2.034)	0.639	1.083 (0.586 – 2.001)	0.799	1.064 (0.577 – 1.960)	0.843
40代ダミー	1.088 (0.619 – 1.911)	0.769	1.078 (0.617 – 1.884)	0.793	1.072 (0.613 – 1.875)	0.808	1.049 (0.576 – 1.909)	0.876	1.041 (0.574 – 1.885)	0.896
50代ダミー	1.131 (0.622 – 2.057)	0.687	1.125 (0.620 – 2.040)	0.699	1.113 (0.614 – 2.017)	0.725	1.114 (0.588 – 2.111)	0.74	1.103 (0.585 – 2.077)	0.763
50代ダミー	0.924 (0.437 – 1.952)	0.836	0.906 (0.429 – 1.911)	0.795	0.957 (0.453 – 2.021)	0.907	0.912 (0.407 – 2.043)	0.823	0.954 (0.428 – 2.126)	0.908
70代以上ダミー	1.896 (0.606 – 5.934)	0.272	1.838 (0.588 – 5.746)	0.295	1.766 (0.568 – 5.496)	0.326	2.078 (0.606 – 7.128)	0.245	2.019 (0.594 – 6.861)	0.26
個人年収（参照群：4-6百万）										
0円ダミー	0.615 (0.355 – 1.065)	0.083	0.615 (0.355 – 1.063)	0.081	0.580 (0.336 – 1.001)	0.051	0.545 * (0.302 – 0.985)	**0.044**	0.513 * (0.284 – 0.927)	**0.027**
1-2百万ダミー	0.940 (0.602 – 1.470)	0.788	0.958 (0.614 – 1.495)	0.850	0.919 (0.590 – 1.431)	0.708	0.864 (0.535 – 1.397)	0.552	0.828 (0.513 – 1.336)	0.439
2-4百万ダミー	0.909 (0.584 – 1.414)	0.671	0.910 (0.586 – 1.414)	0.675	0.881 (0.568 – 1.368)	0.574	0.825 (0.512 – 1.327)	0.427	0.797 (0.496 – 1.282)	0.349
6-8百万ダミー	1.200 (0.716 – 2.009)	0.489	1.199 (0.716 – 2.009)	0.490	1.184 (0.708 – 1.979)	0.519	1.050 (0.600 – 1.835)	0.865	1.029 (0.590 – 1.793)	0.921
8-10百万ダミー	1.022 (0.497 – 2.102)	0.952	0.999 (0.485 – 2.056)	0.997	0.973 (0.473 – 2.000)	0.941	0.960 (0.439 – 2.099)	0.919	0.942 (0.431 – 2.055)	0.88
10百万以上ダミー	0.418 * (0.200 – 0.873)	**0.02**	0.417 * (0.200 – 0.871)	**0.020**	0.411 * (0.198 – 0.856)	**0.018**	0.337 ** (0.152 – 0.748)	**0.007**	0.330 ** (0.149 – 0.729)	**0.006**
不明・非回答ダミー	0.695 (0.427 – 1.132)	0.144	0.704 (0.433 – 1.147)	0.158	0.694 (0.428 – 1.126)	0.139	0.644 (0.381 – 1.089)	0.101	0.631 (0.374 – 1.066)	0.085
居住地域（参照群：関東）										
北海道ダミー	1.915 (0.959 – 3.826)	0.066	1.893 (0.948 – 3.781)	0.070	1.994 * (1.002 – 3.969)	**0.049**	2.034 (0.964 – 4.291)	0.062	2.143 * (1.015 – 4.526)	**0.046**
東北ダミー	1.222 (0.713 – 2.095)	0.465	1.231 (0.718 – 2.108)	0.450	1.242 (0.727 – 2.120)	0.428	1.204 (0.672 – 2.156)	0.532	1.213 (0.679 – 2.168)	0.514
中部ダミー	1.548 * (1.074 – 2.230)	**0.019**	1.511 * (1.051 – 2.171)	**0.026**	1.546 * (1.076 – 2.220)	**0.018**	1.542 * (1.041 – 2.286)	**0.031**	1.575 * (1.064 – 2.333)	**0.023**
近畿ダミー	1.149 (0.802 – 1.646)	0.448	1.144 (0.798 – 1.639)	0.464	1.160 (0.811 – 1.659)	0.416	1.100 (0.746 – 1.622)	0.631	1.115 (0.757 – 1.643)	0.581
中国ダミー	1.251 (0.709 – 2.208)	0.439	1.235 (0.700 – 2.178)	0.465	1.278 (0.727 – 2.248)	0.394	1.245 (0.674 – 2.300)	0.485	1.279 (0.694 – 2.357)	0.43
四国ダミー	0.860 (0.382 – 1.936)	0.715	0.872 (0.387 – 1.968)	0.742	0.911 (0.405 – 2.052)	0.822	0.801 (0.332 – 1.933)	0.621	0.834 (0.347 – 2.009)	0.686
九州ダミー	2.002 ** (1.214 – 3.299)	**0.006**	2.013 ** (1.220 – 3.319)	**0.006**	1.981 ** (1.204 – 3.258)	**0.007**	2.067 ** (1.202 – 3.554)	**0.009**	2.053 ** (1.196 – 3.524)	**0.009**
結婚（参照群：未婚）										
既婚ダミー	0.704 (0.474 – 1.046)	0.082	0.707 (0.476 – 1.048)	0.084	0.696 (0.469 – 1.031)	0.071	0.650 * (0.423 – 0.997)	**0.048**	0.641 * (0.418 – 0.982)	**0.041**
不明・非回答ダミー	1.285 (0.441 – 3.751)	0.646	1.255 (0.427 – 3.689)	0.679	1.254 (0.432 – 3.639)	0.677	0.961 (0.299 – 3.084)	0.946	0.950 (0.296 – 3.052)	0.931
子ども（参照群：子なし）										
子ありダミー	1.701 * (1.133 – 2.553)	**0.01**	1.699 * (1.134 – 2.546)	**0.010**	1.755 ** (1.171 – 2.630)	**0.006**	1.935 ** (1.247 – 3.003)	**0.003**	1.987 ** (1.282 – 3.080)	**0.002**
不明・非回答ダミー	0.784 (0.271 – 2.267)	0.653	0.793 (0.273 – 2.308)	0.670	0.803 (0.279 – 2.313)	0.685	1.134 (0.357 – 3.605)	0.832	1.164 (0.366 – 3.703)	0.797
Random Effects										
σ2	3.29		3.29		3.29		3.29		3.29	
τ00（個人レベル）	4.53		4.54		4.49		5.33		5.29	
ICC	0.58		0.58		0.58		0.62		0.62	
N	1138		1138		1138		1138		1138	
Observations	11380		11380		11380		10242		10242	
固定 R2 / 固定・変量 R2	0.062 / 0.605		0.061 / 0.606		0.067 / 0.605		0.063 / 0.642		0.069 / 0.643	
AIC	328221.925		328217.837		328212.246		281790.222		281788.293	

とは，AIに比べて人間に対する向社会性が高いことが示されている。

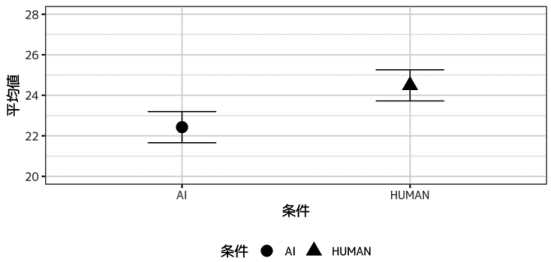

図２　対AI条件および対人間条件におけるSVOスライダー得点

3.3.2　信頼に関する比較

はじめに，４つの項目からなる修正した一般的信頼尺度の妥当性を検証するために対人条件とAI条件の一般的信頼項目についてα係数を算出したところ，対人一般的信頼については$\alpha=0.89$，対AI一般的信頼については$\alpha=0.87$であった(3)。したがって，このいずれの質問項目についても一貫性があるとみなして，その合計得点をもとに分析する。

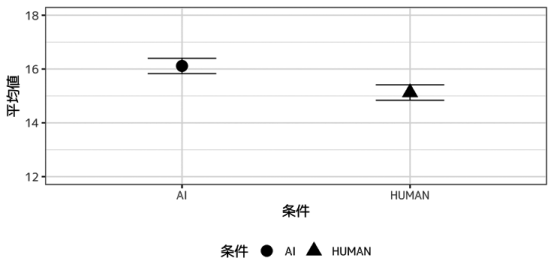

図３　対AI条件および対人間条件における信頼

図３では修正した一般的信頼尺度の各項目の合計の平均値について対AI条件と対人間条件の95%信頼区間を示している。この２群において対応のあるt検定を行ったところ，$t(1137)=7.07$, $p<.001$でAI条件の方が高いことが明らかとなった。

3.3.3　対他的親密性に関する比較

図４ではIOS尺度の平均値について，対AI条件と対人間条件の95%信頼区間を示している。この

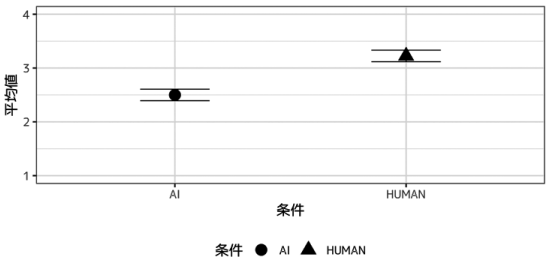

図４　対AI条件および対人間条件におけるIOS得点

２群において対応のあるt検定を行ったところ，$t(1137)=14.203$，$p<.001$で人間条件の方が高いことが明らかとなった。このことは，AIに比べて人間に対して親近感を抱いていることを示している。

3.3.4　小括

本項においては調査項目の中でも信頼，向社会性を反映する社会的価値指向性，共感性を示す対他的親密性について，対人条件と対AI条件を比較した。その結果，人間に比べてAIの方が信頼される一方で，AIに比べて人間に対して向社会的であり，共感することが明らかとなった。

3.4　対他的親密性の変化

最後に，AIとの共に働く作業が親近感に与える影響を分析するために，公共財ゲーム前に実施した一般的なAIに対するIOS尺度と，公共財ゲーム以降に実施した，同じゲームに参加した各仮想エージェントを対象としたIOS尺度の平均値の比較を行う。これにより，検討課題４について明らかにする。

表３には公共財ゲーム実験前のAIに対するIOS尺度と公共財ゲーム実験後の仮想エージェントAおよびBに対するIOS尺度のスコアを応答変数として，仮想エージェントAおよびBのダミー変数を説明変数とした分析結果を示している。コントロール群は一般的なAIに対するIOS得点である。Model６では公共財ゲーム実験前の一般的なAIに対する共感性に比べて，いずれの仮想エージェン

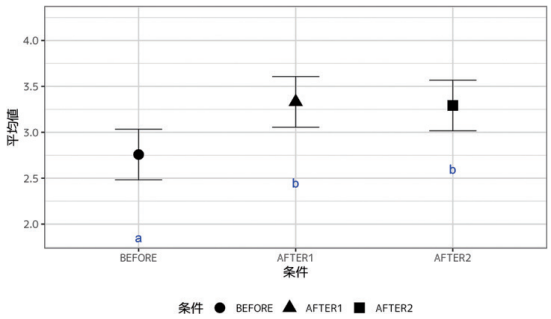

図５　公共財ゲーム前後のIOS得点

表３　分析結果

	AIエージェントに対するIOS尺度			
	Model 6		Model 7	
Predictors	*Estimates*	*p*	*Estimates*	*p*
(Intercept)	2.444 *** (2.338 – 2.550)	**<0.001**	2.981 *** (2.075 – 3.887)	**<0.001**
仮想エージェントA	0.574 *** (0.465 – 0.682)	**<0.001**	0.574 *** (0.466 – 0.682)	**<0.001**
仮想エージェントB	0.534 *** (0.426 – 0.643)	**<0.001**	0.534 *** (0.426 – 0.642)	**<0.001**
認知反射テスト			-0.062 * (-0.124 – -0.001)	**0.045**
			社会経済的要因を統制済み	
Random Effects				
σ2	1.75		1.72	
τ00（個人レベル）	1.16		1.14	
ICC	0.4		0.4	
N	906		906	
Observations	3414		3414	
固定 R2 / 固定・変量 R2	0.023 / 0.412		0.051 / 0.430	
AIC	12610.957		12645.425	

トに対しても公共財ゲーム以降に改善していることが示されている。図５はModel６にもとづき，条件間の最小二乗平均値と95%信頼区間ならびにTukeyの補正を行った多重比較の結果をCLD（Compact Letter Display）によって示している。同じ記号で表されたもの同士において差があるとは言えないということを示している。その結果，公共財ゲーム実験以前に比べて，以降において仮想エージェントAおよびBに対する親近感が改善していることが示されている。

Model７においては認知反射テストおよびその他の社会経済的要因を統制した分析結果を示している。この結果はModel６で得られた結果が頑健であること，認知反射テストのスコアが高さと共感性に負の相関があることを示している。この結果は，人間とAIと共に働くことで，人間のAIに対する親近感が改善し得ることを示唆している。

4　考察

4.1　まとめ

本研究の概要は以下の通りである。

- 人間はAIと協力できる。協力的仮想エージェントが多い方が，人間も公共財ゲームで貢献する。N–1期目におけるAIエージェントの貢献額の平均値がN期における貢献額に影響を与える。
- AIと協力しやすい人としにくい人は存在する。公共財ゲームにおける貢献額に対して，対人信頼および対人SVOは影響を及ぼさないが，対AISVOは影響を及ぼしている。
- 人間のAIに対する認識は，人間に対する信頼に比べて，AIに対する信頼が高い。しかし，向社会性と共感性についてはAIに比べて，人間に対して高い。
- AIと共に作業を行うことで，AIに対する共感性が高まる可能性がある。

4.2　ディスカッション

本研究から以下の点が明らかになった。

検討課題１について，人間とAIは協力可能であるが，協力的仮想エージェントが多い条件の方が，N–1期目におけるAIエージェントの貢献額が多い方が人間は公共財ゲームで協力することが明らかとなった。この結果は，AIエージェントが協力的であると，人間はより協力することを示唆している。あわせてN–1期目における仮想エージェントと貢献額とN期における人間の貢献額が正の相関をするという一種の条件つき協力（Fischbacher, et. al, 2001）に類似した行動が観察された。

また，公共財ゲームにおける実験参加者の行動を確認すると，貢献額が期を経るほど減少し，チェック問題の正答数が多いほど，すなわち公共財ゲームをしっかり理解しているほど貢献額は減少するという結果が得られている。この結果は，標準的な公共財ゲームにおける先行研究と同様の結果であり，妥当な結果であるといえる。

人間が，どのような仮想エージェントに対しても平均的に協力しないことはないことは，人間とAIが協力・共生できる可能性を示唆している。特に，協力的エージェントが多いと協力すること，N–1期目における仮想エージェントの貢献額とN期における人間の貢献額の正の相関をすることを踏まえると，少なくとも当初は，AIエージェントが協力するように設計することで人間の協力行動を引き出すことができると考えられる。ただし，今回はあくまでも単純な公共財ゲームを用いている。例えば処罰や報酬などの仕組み（Balliet et. al, 2011）を導入した公共財ゲームにより複雑な状況についても検討する必要があるであろう。

検討課題2について，AIと協力しやすい人としにくい人は存在するが，その弁別においてSVOの応用が可能であることが明らかとなった。公共財ゲームにおいて，SVOが影響を及ぼすことはその定義および先行研究を鑑みても明らかであるが（De Cremer, and Van Lange, 2001；De Cremer and Van Dijk, 2002），本研究ではAIについてもSVOの観点から分析可能であることが示された。

検討課題3について，AIに対する信頼が高いものの，AIに対する向社会性ならびに共感性が低いことが明らかとなった。本研究ではこのような結果が出た原因の解明まで至らないが，2つの原因が推測される。1つは実験参加者がクラウドソーシングワーカーであることにある。比較的新しい技術に親和性が高い可能性がある。そしてもう1つの原因の可能性としては，少しずつAIが社会に浸透してきた結果として，AIに対する信頼が高まっている可能性もある。ただし，笠木（2018）によれば，人間に対する信頼は交流の中で徐々に上昇していく一方で，機械に対する信頼は，最初は高いものの，徐々に低下することもあると指摘されている。したがって，一種のバイアスとしてAIに対する過剰な信頼，ないしは期待が反映されている可能性があることに留意する必要がある。

検討課題4について，仮想エージェントに対する共感性は，作業後に高まることが明らかとなった。すなわち，AIエージェントと人間が共に作業を行うことで，AIエージェントに対して共感性が低い個人であったとしても，その程度が改善し得ることを示唆している。

4.3　今後の課題

今後の課題として，以下の7点をあげる。

第1に，本研究においては，類型化されたAIエージェントを対象とした協力行動の可能性について検討することを目的としているために，デセプションを用いる必要があった。しかしながら，昨今注目を浴びている，人工知能研究所OpenAIが開発しているGPTやGoogleによるBardに代表される生成型AIなどを用いた実験は可能であろう。また，特定のAIとの協力行動については，別の課題にて検討する必要があるであろうし，そのような高度なAIがどのような学習をするのか，その学習モデルに対する社会的認知がどのように影響するのかについても検討する必要があるであろう。

第2に，置かれた状況によるAIエージェントとの行動・認知の差異である。例えば，人生の岐路ともいうべき状況と娯楽状況においてはAIとの行動やAIに対する認知は異なるものになると考えられる。この点については本研究のような実験ゲームではアプローチが難しい可能性がある。

第3に，人間がAIに対して意志や意図を見出すか否かである。我々は何をもって人間の行動であると判断して，AIの行動であると判断しているのであろうか。例えば，コンピュータを介した文字情報や数字情報のみによるコミュニケーションでは，両者を区別できない可能性がある。もし両者を区別できるのであれば，人間はどのように両者を区別しているのであろうか。これについては対戦相手が人間であるかAIであるかわからな

いような状況で繰り返しゲームを行った上で，相手がどのようなエージェントであると予想したか分析することにより，解明できる可能性がある。

第４に，AIの行動が急に変化した時の人間の行動である。今回は，ランダム条件と一定のレンジの間で意思決定を下すように設定した３条件の計４条件による実験を行った。しかしながら，例えばあまり協力的ではなかったAIが協力するようになった時，もしくは協力的だったAIが協力しなくなった時に個人のAIに対する認識はどのように変化するのであろうか。処罰の影響を含めて検討する必要があるであろう。

第５に，AIに対する「信頼」が壊れた時に人間はAIに対してどのような認識をするのであろうか。例えば，ずっと協力し続けてきて，一定の協力に対する信頼が形成されているAIが急に協力しなくなった時には，そのような信頼が破壊されたといえる。信頼が破壊された後には，再び協力を続けたとしても，簡単には信頼が回復されることはないであろう。この時の信頼回復はどのようなプロセスによってなされるのであろうか。これらの点の解明についてはAI倫理に関する意識調査を含めて分析することで（Ikkatai et al, 2022），そのような変化を詳細に分析できると考えられる。

第６に，AIとの付き合い方が苦手な個人に対する介入の方法である。もちろん，AIとの付き合いが苦手な個人がいることは仕方ない。しかしながら，経済・社会の発展や効率化という観点からはAIと付き合っていかないといけない社会の到来は間近であろう。もちろん，AIと関わらないという個人の自由を担保することは必要であるが，AIの受容を進めるような介入のあり方も検討する必要があるであろう。

第７に，AIとの付き合い方に関する文化差の存在に関する検討である。例えば総務省（2016）によれば，仕事のパートナーとしてのAIに対する抵抗感について，日米を比較すると，日本では「上司」がAIであることに対する抵抗感が強い一方で，アメリカでは「同僚」や「部下」がAIであることに対する抵抗感が強いという。このようにAIの受容には文化差が存在し得る。一方で，効率的なAIの利用は経済・社会の発展に必要不可欠であろう。これからの社会においてAIがさまざまな場面で用いられる時代が眼前に迫っている。このような事実を鑑みると，AIに対する意識の文化差が今後の社会・経済の発展にも大きな影響を与える可能性もある。

もちろん，AIを受容しないという自由もあるかもしれない。しかしながら，現在の社会の状況を鑑みると，AIの非受容は社会的・経済的な不利益を多く被る可能性もあり，今後の社会・経済発展のために市民に対するAIの受容に向けた適切な介入手段は是非を含めて検討する余地がある。

例えば，単純接触効果（ザイオンス効果）のように（Zajonc, 1968），日常生活の中でAIに接触したり，AIと作業をする機会が増えるだけで，AIの受容に対する意識が改善する可能性があることを本研究は示唆しているが，より適切な介入についても検討する必要があるであろう。

これからの時代では，AIと協力した作業が求められる場面が増えるであろう。一方，AIとの協力が向いている人もいれば，向いていない人もいると考えられる。どのような人がAIとの協力できるのか，そして向いていない人もどのようにすれば必要に応じて協力できるようになるのか検討していく必要があるであろう。

謝辞

本研究はJSPS科研費19K20634，19H01470，21KK0027および22K18153の助成により実施した。ここに記して感謝申し上げる。

注

（１）人間の代理として行動を行う人工知能をAIエージェントと表す。ゲームのプレイヤー

（意思決定主体）としての人工知能に着目するためにこのような表現を用いている。

（2）一般的信頼に関わる調査項目は以下のとおりである。(a) ほとんどの（人/人工知能エージェント)は基本的に正直である。(b) ほとんどの（人/人工知能エージェント）は信頼できる。(c) ほとんどの（人/人工知能エージェント）は基本的に善良で親切である。(d) ほとんどの（人/人工知能エージェント）は他人を信頼している。(e) 私は，（人/人工知能エージェント）を信頼するほうである。この質問項目の中から，(d) について除いて分析を行った。

（3）なお，5項目による一般的信頼尺度の妥当性を検証した場合，対人条件とAI条件のα係数についてはそれぞれ0.91, 0.88であった。

参考文献

Andreoni, J. (1995). 17. Cooperation in public-goods experiments: kindness or confusion?. Market Failure or Success, 326.

Aron, A., Aron, E.N., & Smollan, D. (1992). Inclusion of other in the self scale and the structure of interpersonal closeness. Journal of personality and social psychology, 63(4), 596.

Balliet, D., Mulder, L.B., & Van Lange, P.A. (2011). Reward, punishment, and cooperation: a meta-analysis. Psychological Bulletin, 137(4), 594-615.

Blount, S. (1995). When social outcomes aren't fair: The effect of causal attributions on preferences. Organizational behavior and human decision processes, 63(2), 131-144.

Burton-Chellew, M.N., El Mouden, C., & West, S.A. (2016). Conditional cooperation and confusion in public-goods experiments. Proceedings of the National Academy of Sciences, 113(5), 1291-1296.

Chen, D.L., Schonger, M., & Wickens, C. (2016). oTree—An open-source platform for laboratory, online, and field experiments. Journal of Behavioral and Experimental Finance, 9, 88-97.

Cialdini, R.B., Brown, S.L., Lewis, B.P., Luce, C., & Neuberg, S.L. (1997). Reinterpreting the empathy–altruism relationship: When one into one equals oneness. Journal of personality and social psychology, 73(3), 481.

Crandall, J.W., Oudah, M., Ishowo-Oloko, F., Abdallah, S., Bonnefon, J.F., Cebrian, M., Shariff, A., Goodrich, M,A., & Rahwan, I. (2018). Cooperating with machines. Nature communications, 9(1), 1-12.

Dafoe, A., Bachrach, Y., Hadfield, G., Horvitz, E., Larson, K., & Graepel, T. (2021). Cooperative AI: machines must learn to find common ground. Nature, 593(7857), 33-36.

De Cremer, D., & Van Lange, P.A. (2001). Why prosocials exhibit greater cooperation than proselfs: The roles of social responsibility and reciprocity. European Journal of personality, 15 (1_suppl), S5-S18.

De Cremer, D., & Van Dijk, E. (2002). Reactions to group success and failure as a function of identification level: A test of the goal-transformation hypothesis in social dilemmas. Journal of Experimental Social Psychology, 38 (5), 435-442.

Eckel, C.C., & Wilson, R.K. (2006). Internet cautions: Experimental games with internet partners. Experimental Economics, 9(1), 53-66.

Fischbacher, U., Gächter, S., & Fehr, E. (2001). Are people conditionally cooperative?

Evidence from a public goods experiment. Economics letters, 71(3), 397-404.

Ferraro, P.J., & Vossler, C.A. (2010). The source and significance of confusion in public goods experiments. The BE Journal of Economic Analysis & Policy, 10(1).

Frederick, S. (2005). Cognitive reflection and decision making. Journal of Economic perspectives, 19(4), 25-42.

Funk, C., Tyson, A., Kennedy, B., & Johnson, C. (2020). Science and scientists held in high esteem across global publics. Pew research center, 29.

後藤晶. (2021). ビッグデータ時代の経済ゲーム実験：クラウドソーシングを用いた大規模公共財ゲーム実験の実施. 情報処理学会論文誌, 62(5), 1246-1260.

Houser, D., & Kurzban, R. (2002). Revisiting kindness and confusion in public goods experiments. American Economic Review, 92(4), 1062-1069.

Ikkatai, Y., Hartwig, T., Takanashi, N., & Yokoyama, H.M. (2022). Octagon measurement: Public attitudes toward AI ethics. International Journal of Human–Computer Interaction, 1-18.

笠木雅史. (2018). 機械・ロボットに対する信頼. 小山虎（編著）信頼を考える：リヴァイアサンから人工知能まで, 勁草書房, pp.225-252.

Longoni, C., Bonezzi, A., & Morewedge, C.K. (2019). Resistance to medical artificial intelligence. Journal of Consumer Research, 46(4), 629-650.

MS&ADインターリスク総研株式会社(2021)「AI（人工知能）を活用したサービス等の受容度調査現在のAIに対する消費者のイメージは３～４年前に比べてポジティブに」, <https://www.irric.co.jp/topics/press/2021/0205.php> Accessed 2022, July 21

Murphy, R.O., Ackermann, K.A., & Handgraaf, M. (2011). Measuring social value orientation. Judgment and Decision making, 6(8), 771-781.

森久美子. (2015). 社会的価値志向性研究の現在：測定法をめぐる問題. 関西学院大学社会学部紀要, (120), 33-51.

文部科学省（2020）「信頼されるAI」, <https://www.mext.go.jp/b_menu/houdou/2020/mext_00487.html> Accessed 2023, March 1.

内閣府（2016）「第５期科学技術基本計画」, <https://www8.cao.go.jp/cstp/kihonkeikaku/index5.html> Accessed 2022, August 16.

OECD,（訳）齋藤長行（2021）「OECD 人工知能（AI）白書：先端テクノロジーによる経済・社会的影響」, 明石書店, 262p.

奥山尚子, 澤田康幸, 八下田聖峰, (2022). 人間か機械か―経済実験による信頼と信頼性, 佐藤嘉倫, 稲葉陽二, 藤原佳典（編著）AIはどのように社会を変えるか―ソーシャル・キャピタルと格差の視点から, 東京大学出版会, pp.91-115.

Paeng, E., Wu, J., & Boerkoel, J. (2016). Human-robot trust and cooperation through a game theoretic framework. Proceedings of the AAAI Conference on Artificial Intelligence 30(1).

消費者庁(2020)「第１回消費者意識調査結果(AIに対するイメージについて)」, <https://www.caa.go.jp/policies/policy/consumer_policy/meeting_materials/assets/consumer_policy_cms101_20316_03.pdf> Accessed 2022, July 21.

総務省（2016）「情報通信白書」, <https://www.soumu.go.jp/johotsusintokei/whitepaper/h28.html> Accessed 2022, July 21.

Tatsukawa, K., Takahashi, H., Yoshikawa, Y., & Ishiguro, H. (2019). Android pretending to have similar traits of imagination as humans

evokes stronger perceived capacity to feel. Frontiers in Robotics and AI, 6, 88.

WIRED, (2017)「「AlphaGo」という“神”の引退と，人類最強の19歳が見せた涙の意味：現地レポート」, https://wired.jp/2017/05/28/future-of-go-summit-day5/, Accessed 2022, August 16.

山岸俊男. (1998). 信頼の構造―こころと社会の進化ゲーム―東京大学出版会, 224p.

Yamagishi, T., Mifune, N., Li, Y., Shinada, M., Hashimoto, H., Horita, Y., Miura, A., Inukai, K., Tanida, S., Kiyonari. T., Takagishi. H., & Simunovic, D. (2013). Is behavioral pro-sociality game-specific? Pro-social preference and expectations of pro-sociality. Organizational Behavior and Human Decision Processes, 120(2), 260-271.

横井良典, & 中谷内一也. (2018). 治療方針の共有が人工知能への信頼に及ぼす影響. 社会心理学研究, 34(1), 16-25.

Zajonc, R.B. (1968). Attitudinal effects of mere exposure. Journal of personality and social psychology, 9 (2p2), 1-67.

Zelmer, J. (2003). Linear public goods experiments: A meta-analysis. Experimental Economics, 6, 299-310.