

エージェントはタスク難易度によって人間の共感を促進する

Agents promote human empathy by task difficulty

津村賢宏^{1,2 *} 山田誠二^{2,1}
Takahiro TSUMURA^{1,2} Seiji YAMADA^{2,1}

¹ 総合研究大学院大学

¹ The Graduate University for Advanced Studies

² 国立情報学研究所

² National Institute of Informatics

Abstract: 人間から共感を引き出す擬人化エージェントを実現する方法として、エージェントと人間間のタスクに注目し、タスク難易度とタスク内容が人間の共感を促進する可能性について、仮説を立てて実験的に調査した。実験は3要因混合計画で、タスク難易度とタスク内容とタスク前後の共感値の $2 \times 2 \times 2$ の8条件である。結果、タスク内容要因の主効果は認められず、タスク難易度が高いとき、タスク難易度が低いときよりも共感を促進した。

1 はじめに

人間は社会でタスクをする上で、多種多様な道具を用いて生活している。メディア・イクエーションとは、人間が人工物を人間のように扱うことである [1]。これにより、人間は人工物に対しても類似の感情を抱くことが示された。実際に、人間が人間に対して共感すると同様に、人工物に対して共感をしている例がある。我々が共感している人工物の代表例として、掃除ロボットやペット型ロボット、対戦ゲームのキャラクター、オンラインショッピングやヘルプデスクでサービスを行う擬人化エージェントなどがある。これらはすでに人間社会で用いられており、人間と共存している。また、擬人化エージェントの外見は用途や使用環境によって様々である。このように、人間と擬人化エージェントはすでに競争タスクや協調タスクにおいてインタラクションを行っている。一方で、人間の中にはエージェントを受け入れることができない人も一定数存在している [2][3][4]。

これからの人間社会にエージェントが浸透していくためには道具として便利であることとは別に、人間に受け入れられやすい要因をエージェントが持っている必要がある。そこで、人間がエージェントとの関係を良好にしていく方法の1つとして、人間からエージェントに対して共感を抱いてもらうことを考えた。共感することによって、人間はエージェントに対して積極的な行動をとり、エージェントを受け入れやすくなる。共感を引き起こす要因として、言語情報や非言語情報、

状況や関係など様々な研究が行われてきたが、我々はタスクに注目し、タスク難易度とタスク内容が共感にどのような影響を与えるのかを実験的に検証する。

本研究では、人間とエージェントのタスクに注目し、タスク難易度とタスク内容によって人間がエージェントにより共感するかを調査するための実験を行う。

2 関連研究

心理学分野ではこれまで共感を注目し研究されてきた。Omdahl[5] は共感を大きく3種類に分類し、(1) 他者の感情状態に対する感情的反応である感情的共感 (affective empathy) と (2) 他者の感情状態の認知的理解である認知的共感 (cognitive empathy) と (3) 上記の2つを含んだ共感と定義した。Preston と De Waal[6] は共感的な反応の中心には、観察者が対象の主観的な感情状態にアクセスできるようにするメカニズムが存在することを提案した。彼らによって Perception-Action Model (PAM) が定義されて、共感の相違点が統一された。彼らは共感を (a) 他人の感情状態を共有する、またはその影響を受ける、(b) 感情状態の理由を評価する、(c) 他者の視点を特定して取り入れる能力の計3種類として定義した。Olderbakら [7] は、共感の感情特異性に関する理論的・実証的な支持を述べ、6つの基本感情に対する感情的・認知的共感を評価する感情特有の共感アンケートを作成した。

心理学分野では共感の尺度として様々なアンケートが使用されるが、我々は検討の結果、1つのアンケートについて実験で使用することに決めた。Interpersonal

*連絡先：総合研究大学院大学
神奈川県三浦郡葉山町
E-mail: takahiro-gs@nii.ac.jp

Reactivity Index (IRI) は心理学の分野でよく利用されており、共感の特性を調査するために使用されている [8]. 本研究では感情移入と認知的共感をそれぞれ調査するために日本語版 Interpersonal Reactivity Index (IRI-J) を実験に適するように変更して使用した [9].

Human-Agent Interaction(HAI) や Human-robot Interaction(HRI), Human-computer Interaction(HCI) の分野では人間とエージェントやロボット間の共感を研究している. HCI の分野において、共感に注目した研究をしている. Wright と McCarthy[10] は HCI における共感を使用した研究を引用しながら共感の利用について議論している. Pratte ら [11] は共感ツールに関する 26 の出版物を分析して、共感ツール設計者のためのフレームワークを開発した.

HRI の分野として以下のような研究がされてきた. Leite ら [12] は長期間、子供と対話するソーシャルロボットの共感モデルを提示し、評価するために、小学校で長期調査を実施した. 子供の社会的存在、エンゲージメント、ソーシャルサポートに対する認識を測定した. Zhi ら [13] はロボットが社会的影響を活用することで、近くの傍観者が人間の虐待から介入し、積極的に防御するように誘導できるかどうかを調べた.

また、HAI の分野では以下のような研究がされてきた. Richards ら [14] はユーザがさまざまな言葉による共感の表現にどのような状況で反応するかを理解することで、ユーザの感情や本質的な動機に影響を与えるインテリジェント仮想エージェントを設計するのに重要であると指摘した. Okanda ら [15] は外見に注目し、ロボットに対する友情と道徳性についての日本人の大人の信念が、彼らのアニミズムの傾向と共感に関連して、外観（すなわち、人型、犬のような、卵形）が異なるかどうかを調査した.

協調タスクと競争タスクの研究として、Ruissen と de Bruijn[16] の研究がある. この研究では協調タスクと競争タスクをテトリスを利用して実験された. 結果として、協調タスクでは自他の統合を減少させることはなかったが、競争タスクでは自他の統合を減少させたことを確認した.

また、人間とロボットにおける競争タスクの研究として、Kshirsagar ら [17] の研究がある. 彼らは人間とロボットの同一タスクにおける競争タスクを行った結果、参加者が高性能なロボットよりも低性能なロボットを好むことを示した. Boucher ら [18] は人間とロボットの協調タスクを行い、人間の視線を応用して、ロボットに視線誘導をさせることで、音声による指示より早く人間に目標の対象物を認識させることができた.

タスク難易度の研究として、以下のようなものがある. Fuentes-García ら [19] は難易度の異なるチェス問題解決タスクを用いて、参加者の心拍変動から困難さ、ストレス、複雑さ、認知的欲求などを調査した. Cho[20]

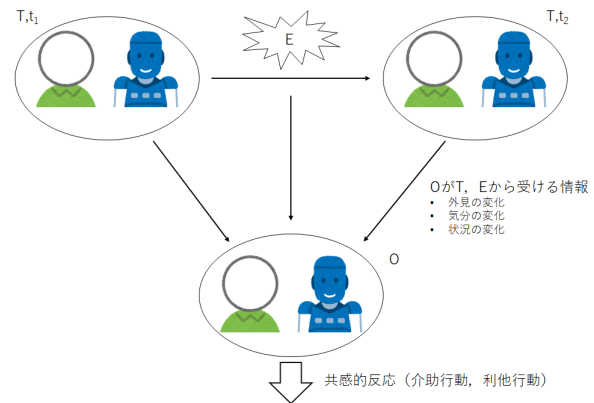


図 1: 共感エージェントの概念図

はインタラクティブシステムの使いやすさと利用頻度を向上させるために、タスク難易度と精神的な作業負荷が必要と考え、人間のまばたきに着目しタスク難易度の自動推定の新しいアプローチを提案した.

Paiva はこれまでの HAI や HRI の研究で扱われているような共感エージェントと言われる人間と共感するエージェントとの関係を定義した. エージェントやロボットと人間との間の共感の定義として、Paiva は共感エージェントを共感されるターゲットとする場合と共感するオブザーバーとする場合の 2 つの異なる方法で表して、それらを図にした [21][22][23].

我々はこれらの図を参考にして図 1 にまとめた. 図 1 の T は target, E は event, O は observer である. t は時間経過を表している. T と E から O へと書かれた矢印は、O が共感するための情報である. この情報は表情や行動、周囲の環境変化などである. O はこれらの情報と T の情報も含めて共感的な反応を示す. これが白色の矢印である. 共感エージェントと人間は T と O のどちらかに分かれる. エージェントが T であるときを共感ターゲットエージェント、O であるときを共感オブザーバーエージェントとする. 本研究では共感ターゲットエージェントを用いて人間の共感を促す.

3 実験方法

3.1 実験目的と実験計画

本研究の目的は、共感エージェントとのインタラクションの中で、エージェントがタスク難易度とタスク内容によって、より人間の共感を引き出すことができるかを調査することである. タスク難易度とタスク内容を共感に関連付け、HAI に応用した研究は我々が初めてである.

以上の目的のために、我々は 2 つの仮説を考えた.

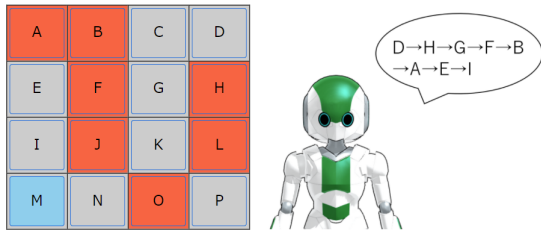


図 2: 高難易度時の共感エージェントとのタスクシーン

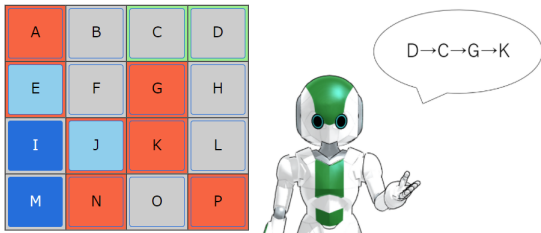


図 3: 低難易度時の共感エージェントとのタスクシーン

- 共感エージェントと競争タスクをする場合、タスク難易度が高いほど人間の共感を抑制する。
- 共感エージェントと協調タスクをする場合、タスク難易度が高いほど人間の共感を促進する。

これらの仮説を調査するために実験を行う。実験は3要因混合計画で行い、その要因として、参加者間要因はタスク難易度とタスク内容の2要因であり、参加者内要因は共感の変動を測るためにタスク前後の共感値を1要因とした。各要因の水準の数は、難易度が2水準（高い、低い）、内容が2水準（競争、協調）、タスク前後の2水準である。全部で8水準ではあるが、参加者内要因が存在するため、参加者は4つの異なる内容の実験から1つだけに参加した。また、従属変数は参加者が抱く共感である。

3.2 実験内容

本実験はオンライン環境で行った。本実験で利用しているオンライン実験はすでに1つの実験方法として利用されるようになっている [24][25][26]。本研究では擬人化エージェントに対する人間の共感を促進させることが目的である。擬人化エージェントとタスクを行う場合、現実ではなくPCを介して利用する環境が想定されるため、オンライン環境であっても同様の効果が得られると考えた。

タスクを行う前に、エージェントに対する共感を測定するためにアンケート調査を行った。この時点では、

タスク難易度やタスク内容について参加者は判断できないようになっている。

タスクは競争と協調の内容によって設定が異なるが、タスク内容が大幅に変わることがないように、共通点として4×4の16マスの範囲内でマスを移動するものにした。すべてのタスクに共通していることは、参加者は必ず左下のマスから移動し、エージェントは必ず右上のマスから移動することである。マス移動は現在地点から上下左右のマスであり、同じマスには各々一度しか通過できない。マス移動は交互に行われるが、エージェントのマス移動は開始時点で表示されている。エージェントが事前にマス移動を提示している理由として、事前に表示をしていない場合、タスクの難易度が高まり、難易度の差が少なくなってしまうからである。エージェントのマス移動まで予想しながら選択することにより、難易度の比較に影響が及ばないように単純なタスクにした。すべてのタスクにはチェックポイントが存在しており、この箇所を通過することが目的になる。参加者が選択したマスを青色、次に移動できるマスを水色、エージェントの選択したマスを薄緑色、チェックポイントを朱色とした。

タスクは合計で3回行われ、すべてマス移動によってチェックポイントを通過することは同様であるが、チェックポイントの位置が異なる。タスクの難易度によって、マス移動の回数が変わり、高難易度では8回、低難易度では4回である。そして、競争タスクでは、チェックポイントを相手より多く通ることを目的とし、協調タスクでは協力してすべてのチェックポイントを通ることが目的となる。その他、詳しい条件に関しては、後の項目で説明する。

また、エージェントの外見や動作も統一しており、本実験で調査したい要因以外の影響を減らしている。参加者はタスク難易度（高い、低い）とタスク内容（競争、協調）を組み合わせたい計4条件のどれか1つの内容でエージェントとインタラクションを行う。

その後、タスク前と同様にエージェントに対して感じた共感をアンケート調査によって集計した。その後、自由記述で実験の感想を書くようお願いした。

3.3 要因の詳細設定

3.3.1 タスク難易度

本実験ではタスク難易度を2種類用意した。これらの難易度については図2と図3の2種類である。難易度の違いとして、以下の条件で行った。

- マス移動を高難易度は8回、低難易度は4回。
- チェックポイントの数が高難易度は7箇所、低難易度は5箇所。

- 競争タスクの場合、高難易度はチェックポイントを4箇所以上通るように行動し、低難易度はチェックポイントを3箇所以上通るように行動する。
- 協調タスクの場合、高難易度はチェックポイントを協力して7箇所すべて通るように行動し、低難易度はチェックポイントを協力して5箇所すべて通るように行動する。

マス移動のための試行回数を半分に減らすことと、参加者とエージェントのチェックポイントの通過を予想を単純化することで、試行回数が難易度に大きく影響するようにした。

3.3.2 タスク内容

本実験ではタスク内容を2種類用意した。タスク環境をできるだけ近いものにするので外部要因を排除し、タスク内容による人間の共感への影響を測るようにした。タスク内容は競争タスクと協調タスクの2種類である。

競争タスクでは、タスクはチェックポイントの取り合いとなり、難易度によって勝利に必要なチェックポイント数に変化はあるが、先に各チェックポイントを通過した方にポイントが付与される。計3回のタスクの勝率は参加者が最適な行動をとった場合でも、1勝1敗1分けになるようにしている。これは、勝率による人間の共感への影響を抑えるために勝率を調整した。

協調タスクでは、タスクは2人でチェックポイントをすべて通過することが目的になり、難易度によってチェックポイント総数に変化する。計3回のタスクの各自がチェックポイントを通過することのできる最大個数の調整をしている。これにより、奇数存在するチェックポイントを常に一方が通過しすぎないようにした。また、両者が同じチェックポイントを通過しても意味はない。

3.4 アンケート

参加者はタスク前後にアンケートに回答した。アンケートは共感の特性を調査するための Interpersonal Reactivity Index (IRI) から、本実験に適するように変更した12項目のアンケートを利用した [8][9]。2回のアンケートは同様のものを用いた。使用したアンケートはどちらも IRI を基にしているため、5ポイントリッカート尺度 (1:あてはまらない, 5:あてはまる) で調査した。利用したアンケートを表1に示す。Q4, Q9, Q10は逆転項目なため、分析する際に点数を反転させている。

3.5 分析方法

分析は3要因混合計画の多変量分散分析を行った。参加者間要因はタスク難易度の2水準とタスク内容の2水準である。参加者内要因はタスク前後の共感値の2水準である。参加者のアンケート結果から、人間の共感を引き出す要因として、タスク難易度やタスク内容がどのように共感の促進に影響を与えるかを調査した。タスク前後で集計した共感の数値を従属変数とする。多変量分散分析には R (R ver.3.6.3) を使用した。

4 実験結果

4.1 実験環境

本実験はクラウドソーシング会社を利用して参加者を募った。参加者には実験参加の報酬として全てのタスクを完了した後に55円を支給した。実験のためにサイトを作成し、PC利用限定で実験を行った。

4.2 参加者

参加者は全部で607人である。しかし、不適当な回答をしている参加者が58人いたため、エラーデータとして排除し、合計549人であった。3要因混合計画の多変量分散分析を行う前に、G*Power 3.1.9.7を用いて3要因多変量分散分析に適した参加者数を計算した。F tests の MANOVA: Repeated measures, within-between interaction を選択して事前分析した。パラメータ設定として、効果量を0.15、有意水準を0.05、検出力を0.80とした。G*Powerの結果、489人の参加者数が適当であるという結果となった。分析に適した人数に揃える作業を行うために、方法として実験への参加順に上から各条件122人を分析対象とした。そのため、分析に使用する参加者は合計で488人である。

年齢は平均で48.21歳 (標準偏差11.04) で、最小が18歳、最大が87歳であった。また、性別は男性が281人で女性が207人であった。

4.3 分析結果

12項目のアンケートについて、全てをまとめた分析を行った。また、共感の分類である感情的共感と認知的共感について分類して分析を行った。多重比較にはホルムの多重比較検定を利用して有意差の存在を検討した。アンケートの分析結果、タスク難易度とタスク前後の2要因の交互作用に有意傾向が認められ、単純主効果を調べた。その単純主効果の分析結果を表2に示す。その結果、タスク聴後のタスク難易度に有意が

表 1: 本実験で使用したアンケートのまとめ

アンケート番号	内容
Q1	もしキャラクターに非常事態が起こった場合、不安で落ち着かなくなる。
Q2	もしキャラクターが感情的になっていた場合、何をしたらいいかわからなくなる。
Q3	もし差し迫った助けが必要なキャラクターを見た場合、混乱してどうしたらいいかわからなくなる。
Q4	もしキャラクターが困っているのを見た場合、気の毒に思わない。
Q5	もしキャラクターが他人にいいように利用されているのを見た場合、その相手を守ってあげたいような気持ちになる。
Q6	キャラクターの話や起こった出来事に心を強く動かされる。
Q7	キャラクターの立場と人間の立場の両方に目を向ける。
Q8	もしキャラクターのことをよく知ろうとした場合、相手からどのように物事がみえているか想像する。
Q9	自分が正しいと思った時に、キャラクターの言い分を聞かない。
Q10	キャラクターの話や起こった出来事に引き込まれることはなく、客観的である。
Q11	キャラクターに起こった出来事が自分の身に起こったらどんな気持ちになるだろうと想像する。
Q12	キャラクターの気持ちに深く入り込んでしまう。

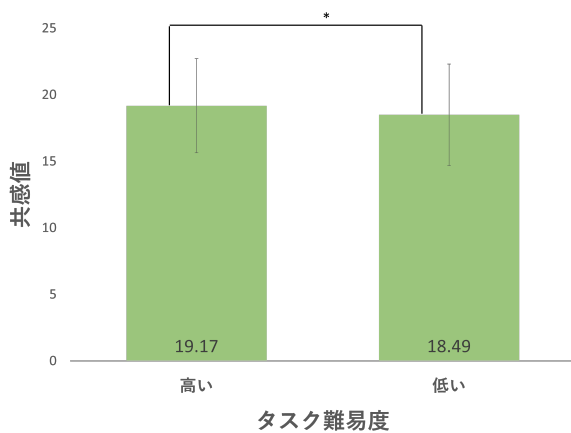


図 4: 共感のタスク後のタスク難易度に対する多重比較の結果

みられたため、図 4 に示す。また、感情的共感の分析結果は交互作用はどこにも認められなかったが、タスク難易度の主効果が認められたため、分析結果を図 5 に示す。

初めに、表 2 より共感の総合の分析結果からタスク難易度とタスク前後の交互作用が認められた [$F(1,484) = 2.793$]。多重比較の結果、タスク後のタスク難易度の単純主効果では、表 2 と図 4 から有意差が認められた (高難易度: 平均 = 37.77, S.D. = 6.040, 低難易度: 平均 = 36.59, S.D. = 7.089)。また、各タスク難易度条件ごとのタスク前後の単純主効果も、表 2 から全てのタスク難易度条件においてタスク前後で有意差が認められた。以上の分析結果を踏まえて、タスク難易度が高いと共感の低下を抑制していることが示唆された。

一方で、タスク内容とタスク難易度とタスク前後の交互作用が認められなかった。タスク内容とタスク難易度の交互作用も認められず、タスク内容とタスク前後の交互作用も認められなかった。

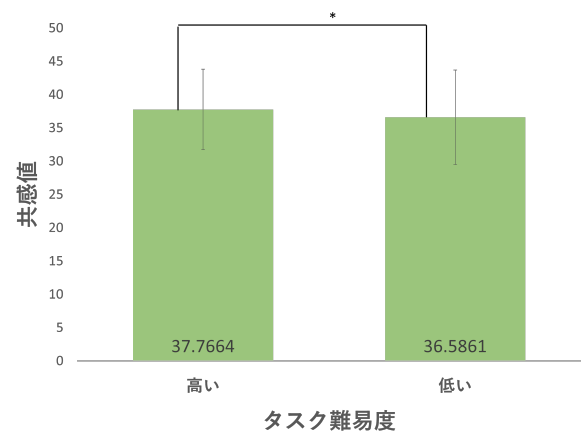


図 5: 感情的共感の主効果の結果

さらに、共感を分類し、感情的共感と認知的共感の 2 種類を分析した結果、認知的共感ではタスク前後の主効果のみ認められ、感情的共感ではタスク前後の主効果の他に、タスク難易度の主効果も認められた [$F(1,484) = 4.915$]。結果として、高難易度では平均 = 19.17, S.D. = 3.539 であったが、低難易度では平均 = 18.49, S.D. = 3.811 であった。このため、難易度が高くなるほど、タスク後の感情的共感の低下を抑制している。

5 議論

本実験では、人間が擬人化エージェントに対して共感を抱く場合に必要となる条件を調査するために行われた。特に、本実験ではタスクに関係する要因を調査することで、エージェントとタスクを行った人間の共感に影響を与える要因の特定を目的とした。この目的のために以下の 2 つの仮説を置き、実験から得られたデータを分析した。

- 共感エージェントと競争タスクをする場合、タス

表 2: タスク難易度とタスク前後の共感の多重比較の分析結果

アンケート	条件	平均	S.D.	要因	<i>F</i>	<i>p</i>	η_p^2
共感 (Q1-12)	高難易度-前	38.56	5.553	タスク前のタスク難易度	1.177	0.2785 <i>ns</i>	0.0024
	高難易度-後	37.77	6.040	タスク後のタスク難易度	3.904	0.0488 *	0.0080
	低難易度-前	38.00	5.857	高難易度のタスク前後	9.852	0.0019 **	0.0391
	低難易度-後	36.59	7.089	低難易度のタスク前後	27.00	0.0000 ***	0.1004

p: **p*<.05 ***p*<.01 ****p*<.001

ク難易度が高いほど人間の共感を抑制する。

- 共感エージェントと協調タスクをする場合、タスク難易度が高いほど人間の共感を促進する。

その結果、2つの仮説は支持されることはなかった。しかし、本実験の結果からタスク内容に関係なく、タスク難易度が高いほど、人間の共感の低下を抑制した。仮説ではタスク内容によって、難易度が人間の共感に影響を与える効果が異なると考えていた。その理由として、Ruissen と de Bruijn[16] の研究から、タスク内容により共感エージェントに対して人間が抱く印象が異なることが示されていたためである。また、協調タスクにおいて、人間は協力する相手によってパフォーマンスを向上させることや好印象を抱くのに対して、競争タスクでは敵対する相手のことを考えることが減るため、上記のような仮説に至った。そして、タスク難易度による共感の変化として、難易度が高くなるほど精神的負荷がかかり、パフォーマンスにも影響が及ぶ。Kshirsagar ら [17] の研究では、競争タスクの中で参加者は対戦相手のロボットが低性能のほうが良いと回答している。これは、ロボットの能力が高いほど、競争の難易度が高くなり、精神的負荷がかかるためと考える。このため、競争タスクにおいて、タスク難易度が高いほど人間の共感は抑制すると考えた。一方で、協調タスクは、協力相手の意図を素早く読み取ることににより、タスクパフォーマンスを向上させていたが、これは認知的共感の分類にある視点取得に関係があると考えられる。相手の立場になって、行動を読み取ることで、タスクを円滑に進めていることから、協調タスクにおいて、タスク難易度が高いほど人間の共感は促進すると考えた。

しかし、本実験ではタスク内容に関係がなく、タスク難易度が高いほど共感の低下を抑制した。この結果を議論する上で、タスク前後の共感の変化に注目する必要がある。本研究において、他の研究ではあまり試されていない手法として、タスクの前後で同様のアンケートを測ることがある。タスク後のみの調査をしていた場合では、タスク難易度が高いほど人間の共感を高めたことになるが、タスク前の調査を行うことで、実際はタスクを通して共感が低下していたことが判明した。この結果はタスク後の変化が必ずしもタスク前より良

い結果になるとは限らないことを示した。そして、本研究で注目した要因のみではタスク後に人間の共感を促進することはできないことが示された。また、本実験においてはタスクによって共感が低下したことから、タスクはエージェントに対する人間の共感を低下させる可能性を示した。ただ、本実験の制限として、タスク難易度とタスク内容以外の要因を排除するために、タスクを単純にしたことでタスク自体を退屈に感じ、共感を低下させた可能性がある。

その一方で、タスク難易度がタスク内容に影響せず、人間の共感の低下を抑制したことは、将来的に共感エージェントが人間社会で共存する際に、人間の共感の制御に効果的な要因になることが考えられる。タスク難易度を適当に設定することで、エージェントに対して必要以上に共感をさせることがなく、適切な距離感を保つことができるようになる。

共感を感情的共感と認知的共感に分類して分析した結果、感情的共感にはタスク前後の主効果とタスク難易度の主効果が認められた。タスク難易度に主効果が認められた理由として、タスク難易度が高くなることで、タスクによる精神的負荷が増加したことによって感情的共感が高まったと考えられる。感情的共感とは他者が経験している・経験しようとする感情状態を感じ取ることで、自身も同じような感情状態になる。タスクによる精神的負荷が両者の感情状態に影響することから、感情的共感だけが高まった。また、認知的共感にはタスク前後の主効果のみが認められ、タスク難易度やタスク内容の主効果は認められなかった。これは認知的共感が他者の思考や感情を自分に置き換えて想像し、相手の視点で想像する必要があるため、本実験のタスクでは認知的共感を高めることができない。

6 まとめ

本研究は擬人化エージェントに対して、人間が共感をする要因と条件の一部として、人間とエージェントのタスクに着目し、タスク難易度とタスク内容を調査した。本研究では2つの仮説を立てて実験を行った。その結果、2つの仮説はどちらも支持されなかった。しかし、実験の結果、人間とエージェントのタスクにおい

て、本実験ではタスク後に共感を低下させることが分析から発見された。この結果は、タスク前後に共感を測るアンケートをとることによって、共感の変化を分析することができたため判明した。そして、タスク内容は共感に対する影響が認められなかったが、タスク難易度は共感に影響を与えることが有意に認められた。分析からタスクの難易度が高いほど、タスク後の共感の低下を抑制することが判明した。また、共感を分類し、感情的共感と認知的共感に分けて分析した場合、感情的共感にはタスク難易度の主効果が認められ、タスク難易度が高いほど感情的共感を高めた。今後の研究として、人間と擬人化エージェントのタスク時に人間が抱く共感が低下する場合が確認できたため、人間の共感に影響を与える要因を調査し、状況に適した人間に共感を抱かれるエージェントの開発が考えられる。

参考文献

- [1] Byron Reeves and Clifford Nass. *The Media Equation: How People Treat Computers, Television, and New Media like Real People and Places*. Cambridge University Press, USA, 1996.
- [2] Tatsuya Nomura, Takayuki Kanda, and Tomohiro Suzuki. Experimental investigation into influence of negative attitudes toward robots on human–robot interaction. *AI & SOCIETY*, 20(2):138–150, Mar 2006.
- [3] T. Nomura, T. Kanda, T. Suzuki, and K. Kato. Prediction of human behavior in human–robot interaction using psychological scales for anxiety and negative attitudes toward robots. *IEEE Transactions on Robotics*, 24(2):442–451, 2008.
- [4] Tatsuya Nomura, Takayuki Kanda, Hiroyoshi Kidokoro, Yoshitaka Suehiro, and Sachie Yamada. Why do children abuse robots? *Interaction Studies*, 17(3):347–369, 2016.
- [5] B. L. Omdahl. *Cognitive appraisal, emotion, and empathy*. Lecture Notes in Computer Science. Psychology Press, New York, 1 edition, 1995.
- [6] Stephanie D. Preston and Frans B. M. de Waal. Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences*, 25(1):1–20, 2002.
- [7] Sally Olderbak, Claudia Sassenrath, Johannes Keller, and Oliver Wilhelm. An emotion-differentiated perspective on empathy with the emotion specific empathy questionnaire. *Frontiers in Psychology*, 5:653, 2014.
- [8] Mark H. Davis and Mark H. Davis. A multi-dimensional approach to individual difference in empathy. In *JSAS Catalog of Selected Documents in Psychology*, page 85, 1980.
- [9] Toshiyuki Himichi, Hidekazu Osanai, Takayuki Goto, Hiroyo Fujita, Yuta Kawamura, Mark H. Davis, and Michio Nomura. Development of a japanese version of the interpersonal reactivity index. *The Japanese Journal of Psychology*, advpub, 2018.
- [10] Peter Wright and John McCarthy. Empathy and experience in hci. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '08, page 637–646, New York, NY, USA, 2008. Association for Computing Machinery.
- [11] Sydney Pratte, Anthony Tang, and Lora Oehlberg. *Evoking Empathy: A Framework for Describing Empathy Tools*. Association for Computing Machinery, New York, NY, USA, 2021.
- [12] Iolanda Leite, Ginevra Castellano, André Pereira, Carlos Martinho, and Ana Paiva. Empathic robots for long-term interaction. *International Journal of Social Robotics*, 2014.
- [13] Xiang Zhi Tan, Marynel Vázquez, Elizabeth J. Carter, Cecilia G. Morales, and Aaron Steinfeld. Inducing bystander interventions during robot abuse with social mechanisms. In *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, HRI '18, page 169–177, New York, NY, USA, 2018. Association for Computing Machinery.
- [14] Deborah Richards, Ayse Aysin Bilgin, and Hedieh Ranjbartabar. Users' perceptions of empathic dialogue cues: A data-driven approach to provide tailored empathy. In *Proceedings of the 18th International Conference on Intelligent Virtual Agents*, IVA '18, page 35–42, New York, NY, USA, 2018. Association for Computing Machinery.
- [15] Mako Okanda, Kosuke Taniguchi, and Shoji Itakura. The role of animism tendencies and empathy in adult evaluations of robot. In *Proceedings of the 7th International Conference on*

- Human-Agent Interaction*, HAI '19, page 51–58, New York, NY, USA, 2019. Association for Computing Machinery.
- [16] Margit I. Ruissen and Ellen R. A. de Bruijn. Competitive game play attenuates self-other integration during joint task performance. *Frontiers in Psychology*, 7, 2016.
- [17] Alap Kshirsagar, Bnaya Dreyfuss, Guy Ishai, Ori Heffetz, and Guy Hoffman. Monetary-incentive competition between humans and robots: Experimental results. In *Proceedings of the 14th ACM/IEEE International Conference on Human-Robot Interaction*, HRI '19, page 95–103. IEEE Press, 2019.
- [18] Jean-David Boucher, Ugo Pattacini, Amelie Le-long, Gerard Bailly, Frederic Elisei, Sascha Fagel, Peter Dominey, and Jocelyne Ventre-Dominey. I reach faster when i see you look: Gaze effects in human–human and human–robot face-to-face cooperation. *Frontiers in Neurorobotics*, 6, 2012.
- [19] Juan P. Fuentes-García, Santos Villafaina, Daniel Collado-Mateo, Ricardo de la Vega, Pedro R. Olivares, and Vicente Javier Clemente-Suárez. Differences between high vs. low performance chess players in heart rate variability during chess problems. *Frontiers in Psychology*, 10, 2019.
- [20] Youngjun Cho. Rethinking eye-blink: Assessing task difficulty through physiological representation of spontaneous blinking. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, CHI '21, New York, NY, USA, 2021. Association for Computing Machinery.
- [21] Ana Paiva, João Dias, Daniel Sobral, Ruth Aylett, Polly Sobreperes, Sarah Woods, Carsten Zoll, and Lynne Hall. Caring for agents and agents that care: Building empathic relations with synthetic agents. *Autonomous Agents and Multiagent Systems, International Joint Conference on*, 1:194–201, 01 2004.
- [22] Ana Paiva. Empathy in social agents. *International Journal of Virtual Reality*, 10(1):1–4, Jan. 2011.
- [23] Ana Paiva, Iolanda Leite, Hana Boukricha, and Ipke Wachsmuth. Empathy in virtual agents and robots: A survey. *ACM Trans. Interact. Intell. Syst.*, 7(3), September 2017.
- [24] R.N Davis. Web-based administration of a personality questionnaire: Comparison with traditional methods. *Behavior Research Methods, Instruments, Computers*, 31:572–577, 1999.
- [25] Matthew J. C. Crump, John V. McDonnell, and Todd M. Gureckis. Evaluating amazon’s mechanical turk as a tool for experimental behavioral research. *PLOS ONE*, 8(3):1–18, 03 2013.
- [26] Kazuo Okamura and Seiji Yamada. Adaptive trust calibration for human-ai collaboration. *PLOS ONE*, 15(2):1–20, 02 2020.