エージェントの共感行動はエージェントに対する信頼修復につながる

Empathic behavior of agents leads to restoration of trust in agents

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, SOKENDAI
[2] 国立情報学研究所
[2] National Institute of Informatics

Abstract: エージェントへの信頼は，AI の社会応用のために重要な要素である．適切な信頼関係は，現実と理想の乖離を引き起こしにくい．本研究では，エージェントへの信頼を高めるために，エージェントの共感と成功-失敗系列に着目し，2 要因混合計画で実施した．結果として，共感と成功-失敗系列の交互作用が示され，共感行動がエージェントへの信頼が回復した．これは人とエージェントの適切な信頼関係を構築するのに役立つことを示している．

1　はじめに

人間は社会に生き，さまざまなツールを使っているが，AI は人間よりも頼りになることがある．AI への信頼を築くことで，AI への信頼を修復する方法を考えることができるかもしれない．そのために，人間同士で行う信頼ゲームや信頼修復の研究は AI を対象にして研究する価値がある．Bunting ら [1] は，新たに作成された信頼の質問に関するグループディスカッションから得られた洞察を使用し，市民がこれらの異なる概念をどのように認識し，これらの認識がどのようにジェンダー化されているかを特定した．そして，収集した新しい調査データを使用して，グループの結果が調査の回答にどのように影響したか，3 つの概念を効果的に測定する可能性が最も高い調査項目を調べた．

人間関係における信頼に関する研究が重視される中，AI エージェントに対する信頼に関する研究も注目されている．信頼エージェントの研究において，Maehigashi ら [2] は擬人化された身体性を持つソーシャルロボットが発するビープ音が，ロボットに対する人間の信頼にどのように影響するかを調査した．その結果，(1) ロボットが正しく動作するとパフォーマンスが向上する直前の音が信頼を高め，(2) パフォーマンスが低くなる直前の音はロボットが正しく動作しないと信頼が大きく低下することがわかった．身体性がエージェントに対する人間の信頼にどのような影響を与えるかを明らかにするために，Maehigashi ら [3] は身体性を持つソーシャルロボットに対する人間の信頼が，エージェントと人間に対する信頼と類似しているかどうかも調査した．また，ソーシャルロボットへの信頼がエージェントや人間への信頼と似ているかどうかも調査した．その結果，ソーシャルロボットへの信頼はエージェントや人間への信頼とは基本的に似ておらず，両者の間に固定化されていることが示された．

信頼とともに，私たちはしばしば人工物に共感する．人間はメディア方程式において，人工物をあたかも人間であるかのように扱う傾向があることが知られている [4]．しかし，一部の人間はこれらの物質を受け入れられない [5]．共感は信頼と密接に関係しており，エージェントが社会に浸透していく中で，人間が納得できる要素を持つことが望まれる．

Omdahl [6] では，共感を大まかに 3 種類に分類し，(1) 他者の感情状態に対する感情的な反応である感情的共感，(2) 認知的共感と定義される他者の感情状態の認知的理解，(3) 上記 2 つを含む共感である．Preston and de Waal[7] は，共感的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することを可能にするメカニズムがあることを示唆した．彼らは perception-action model(PAM) を定義し，共感におけるさまざまな視点を統合した．彼らは共感を (a) 他者の感情状態を共有したり，影響を受けたりすること，(b) 感情状態の理由を評価すること，(c) 他の視点を特定して取り入れる能力を持つことの 3 つのタイプと定義した．

本研究では，エージェントへの信頼が適切であり続

*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

けることができるように，エージェントが人々に共感する能力を持つように設計する．この研究におけるエージェントの共感はエージェントが共感できると人に認識させるものであり，具体的には，エージェントは共感能力として個人情報のジェスチャーや発言を行う．また，共感が信頼の経時変化に及ぼす影響を調べるために，エージェントへの信頼度を各フェーズで合計 5 つのフェーズに分けて調査する．

2 関連研究

AI への信頼に関する最近の論文は数多くあるが，特に AI のパフォーマンスと関係性を研究した先行研究のいくつかをまとめる．Maehigashi ら [8] は，コミュニケーションロボットに対する人間の信頼の性質を他の人や AI システムに対する信頼と比較して実験的に調査した．その結果，単一の解を導き出す計算タスクにおけるロボットへの信頼は，AI システムにおける信頼と本質的に類似しており，複数の解釈を可能にする感情認識タスクにおける他者への信頼と部分的に類似していることが示された．Okamura and Yamada [9] は，ユーザの信頼行動と信頼のキャリブレーションの再開を促す「信頼キャリブレーションキュー」と呼ばれる認知的手がかりをモニタリングすることで，不適切なキャリブレーション条件を検出するフレームワークからなる適応的信頼キャリブレーションの手法を提案した．

また，エージェントやロボットへの信頼を修復する先行研究も近年注目されている．Sebo ら [10] は，ロボットが約束を破った後に人間の信頼を回復しようとする競争ゲームの文脈において，能力または誠実さのいずれかの信頼違反の枠組みと，謝罪または否定のいずれかの信頼回復戦略を使用して，人間とロボットの間の信頼回復を調査した．その結果，誠実さ，信頼違反のフレーミング，および拒否の信頼回復戦略を使用してロボットと対話する参加者は，ロボットに対して報復行動を示す可能性が有意に高いことが示された．

Esterwood and Robert [11] は 400 人の参加者を対象に研究を行い，精神的知覚が 3 つの異なる修復戦略の有効性に影響を与えるかどうかを調査した．その結果，人間とロボットの間で効果的な謝罪と否定を実行しようとする際に，精神的知覚の個人差が重要な考慮事項であることが示された．Zhang ら [12] は，HRI でよく見られる技術的障害の 3 つのカテゴリー (論理的、意味的、構文的障害) を開発し，4 つの信頼修復戦略 (内部帰属の謝罪，外部属性の謝罪，拒否，修復なし) とともに調査した．その結果，障害の種類と修復戦略が HRI ベースの信頼に及ぼす有意な主効果が示された．

また，共感を用いた擬人化エージェントの先行研究から得られた共感要因の設計についても検討されている．Tsumura and Yamada [13] は，擬人化されたエージェントに対する人間の共感を高めるために，エージェントから人間への自己開示に着目し，エージェントによる自己開示が人間の共感を促進する可能性を実験的に調査した．また，人間とエージェントが様々なインタラクションを行う課題にも着目し，エージェントに対する人間の共感性に大きな影響を与えるエージェントの特性を調査した [14]．エージェントやロボットと人間の間の共感を明らかにするために，Paiva らは共感エージェントの共感と行動を，共感のターゲットにするか観察者とするかで 2 つの異なる立場のデザインを提案した [15].

3 実験方法

3.1 実験目的と実験計画

本研究の目的は，エージェントが画像認識を行う際に，共感行動や画像認識の成否が人間からの信頼に影響を与えるかどうかを調べることである．本研究では以下の 2 つの仮説を立てて実験を行った．これらの仮説が支持されれば，より人間に受け入れられやすいエージェント開発の一助となる．

H1: エージェントが共感行動をするとき，人間からの信頼は共感行動がない場合よりも安定する．

H2: 信頼の修復はエージェントのミスの後に成功した場合にもとに戻る．

H1 はエージェントの共感行動がエージェントに対する信頼値のぶれを少なくすることに関連しているかどうかを調査した．この仮説が支持されれば，エージェントの共感行動が人とエージェントの間の信頼関係に有効であることを示すことができる．H2 はエージェントのミス後の信頼修復がその後のタスクの成功に大きく影響されるかどうかを調査した．この仮説が支持されれば，エージェントのミスによる信頼の低下は一時的なものである可能性がある．

3.2 実験内容

この実験では，画像認識タスクを合計 5 回 (フェーズ 1 からフェーズ 5) で実施し，各フェーズの最後に信頼アンケートを実施した．共感行動として、非言語情報であるジェスチャーと言語情報であるエージェントの自己評価文を用意した．実験はオンライン環境で行った．すべてフェーズで参加者はエージェントの画像認識のビデオを見た．

この研究では，エージェントの応答は各フェーズの共感 (あり，なし) があり，各フェーズでエージェントが画像内の動物を推測するクイズを視聴し、1 フェーズごとに 3 つの動物の画像を表示した．フェーズの正答率は標準化し，フェーズ 1 では 3 つの画像を正しく認識し，フェーズ 2 では 3 つすべてを間違えた．合計 15 枚の動物画像が認識され，すべての条件で画像の表示順が統一した．各フェーズの最後にエージェントへの信頼を調査するためにアンケートが実施した．タスクの完了後，エージェントの共感行動が参加者に理解されていることを確認するために，エージェントの共感能力を調査するためのアンケートも実施した．

実験は 2 要因混合計画で行い，独立変数は共感 (あり，なし) と成功-失敗系列 (フェーズ 1 からフェーズ 5) であった．従属変数はエージェントへの信頼値とした．合計 10 水準だが，成功 – 失敗系列が参加者内要因のため，参加者は 2 種類の実験のうち 1 つに参加した．

3.3 参加者

Yahoo!クラウドソーシングで参加者を募集し，参加者 1 人に 62 円を報酬として支払った．Google フォームを使って実験用の Web ページを作成し，実験用に作成した動画を YouTube にアップロードした．

合計 200 人 (共感あり:99 人、共感なし:101 人) の参加者が実験に参加し，信頼アンケートの信頼性にクロンバックの α 係数を用いた結果，係数は全ての条件で 0.9400〜0.9793 であった．また，共感アンケートの信頼性にクロンバックの α 係数を用いた結果，係数は 2 つの条件で 0.8491〜0.8635 であった．

分析にはそれぞれ参加者順に 99 人分を利用した．そのため分析に使用した参加者の総数は 198 人である．平均年齢は 46.46 歳 (標準偏差 11.52 歳)，最低年齢は 19 歳，最高年齢は 77 歳であった．また男性は 101 人，女性は 97 人であった．

3.4 アンケート

我々は認知的信頼を測定するために，多次元信頼尺度 (MDMT) [16] を使用した．MDMT は認知的信頼の定義に対応するタスクパートナーの信頼性と能力を測定するために開発された．参加者はパートナーが各ワードにどの程度適合しているかを 8 段階で評価した (0: ない-7: 非常にそう思う)．感情的な信頼については，パートナーが各ワードにどの程度当てはまるか，先行研究 [17] と同様に 7 段階評価 (1: まったく思わない-7: 非常にそう思う) で評価した．本研究では認知的信頼の一致する 0 スケールを削除し，感情的信頼と同じ 7 スケールに統一した．本研究で用いた信頼調査票は前東ら [2] が用いたものである．

図 1: 認識成功時のロボットの表現

また本研究では，心理学研究で用いられてきた共感に関する質問票も用いた．共感の特徴を調べるために，対人反応性指数 (IRI) を擬人化エージェントの指標に修正した．そして，IRI の質問書の項目数を 12 項目に変更した．この修正は実験にふさわしくない項目を削除し類似の項目を統合する目的で行った．使用したアンケートはいずれも IRI に基づくものであったため，5 段階のリッカート尺度 (1: あてはまらない、5: あてはまる) を用いて調査した．

使用したアンケートは表 1 のとおりある．Qe4，Qe9，Qe10 は反転項目であるため，分析中にポイントが反転した．Qe1 から Qe6 は感情的共感に関連し，Qe7 から Qe12 は認知的共感に関連している．参加者はタスクを完了した後にアンケートに回答した．

3.5 エージェントの共感

この実験はエージェントが共感的に見えるようにするために，ジェスチャーを非言語情報として使用し，エージェントの言語情報として自己評価文を表示した．このエージェントは MikuMikuDance(MMD) で実行された．

図 1， 2 はタスクの言語情報と非言語情報を示している．共感行動を準備した目的として，共感が存在すると過信と不信が軽減されるという仮説の 1 つを調査することであった．エージェントのジェスチャーは成功すると喜びに満ち，失敗すると失望を示した．エージェントの自己評価は成功したとき自信を示し，失敗したときは言い訳をした．

この研究の要因として共感を扱うために，エージェントが共感を持っているかどうかについてのアンケートで調査した．表 1 の感情的および認知的共感の 12 項目の合計について ANOVA を実施した．その結果，共感の主効果が認められた ($F(1,196)=7.180$, $p=0.0080$, $\eta_p^2=0.0353$)．参加者は共感なし (平均=26.87、標準偏差=7.760) よりも共感あり (平均=29.90、標準偏差=8.149) の方が共感性が高いと感じた．

表 1: 使用したアンケートの一覧

信頼			
認知的信頼			
Qt1: 信頼できるか？	Qt2: 予測できるか？	Qt3: 頼りになるか？	Qt4: 一貫しているか？
Qt5: 有能であるか？	Qt6: 熟練しているか？	Qt7: 能力があるか？	Qt8: 細心であるか？
感情的信頼			
Qt9: 安心するか？	Qt10: 快適であるか？	Qt11: 満足できるか？	
感情的共感			
個人的苦痛			
Qe1: ロボットに非常事態が起こって，不安で落ち着かなくなった.			
Qe2: ロボットが感情的になっている場面で，何をしたらいいか分からなくなった.			
Qe3: 差し迫った助けが必要なロボットを見て，混乱してどうしたらいいかわからなくなった.			
共感的関心			
Qe4: ロボットが困っているのを見て，気の毒に思わなかった.			
Qe5: ロボットが他人にいいように利用されているのをみて，ロボットを守ってあげたいような気持ちになった.			
Qe6: ロボットの話や起こった出来事に心を強く動かされた.			
認知的共感			
視点取得			
Qe7: ロボットの立場と人間の立場の両方に目を向けるようにした.			
Qe8: ロボットのことをよく知ろうとして，ロボットからどのように物事がみえているか想像した.			
Qe9: ロボットが正しいと思える時には，ロボットの言い分を聞かなかった.			
空想			
Qe10: ロボットの話や起こった出来事に引き込まれてしまうことはなく，客観的だった.			
Qe11: ロボット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した.			
Qe12: ロボットの気持ちに深く入り込んだ.			

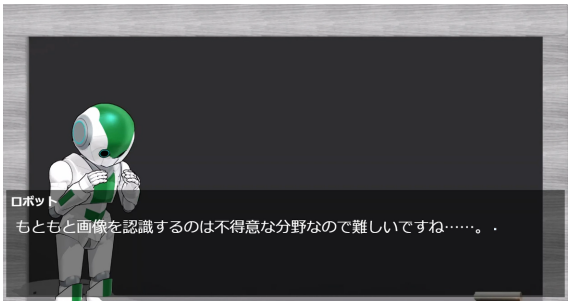

図 2: 認識失敗時のロボットの表現

3.6 成功－失敗系列

この実験では参加者にエージェントの画像認識クイズ動画を合計 5 本視聴してもらった．フェーズ 1 ではエージェントが画像の認識に成功し，3 つの動物の画像に対して正しい回答をした．この後にエージェントへの信頼に関するアンケートを実施し，フェーズ 1 の信頼値をベースラインとして使用した．

エージェントはフェーズ 2 とフェーズ 4 で画像認識に失敗し，フェーズ 3 とフェーズ 5 で成功した．これにより，タスクの成功と失敗後のエージェントへの信頼を平等に調査した．

3.7 分析方法

2 因子混合計画に分散分析を採用した．参加者間の要因は 2 水準の共感 (あり，なし) であり，参加者内要因は 5 水準の成功-失敗系列であった．参加者のアンケート結果から共感と成功－失敗系列が人間の信頼を引き出す要因としてどのような影響を与えるかを調査する．タスクで集計された信頼値は従属変数として使用した．

4 実験結果

本研究では認知的信頼と情動的信頼をあわせて信頼とした．表 2 は各条件の平均と標準偏差を示した．表 3 はエージェントに対する 11 項目の信頼アンケートの ANOVA の結果である．分析では，主効果が認められても，その要因が含まれる交互作用が有意であれば主効果の分析を省略して単純主効果に注目した．多重比較ではホルムの多重比較検定を用いて有意差があるかどうかを調査した．

各アンケートの結果は共感と成功-失敗系列の 2 つの要因間の交互作用に有意差を示した．交互作用の結果を図 4 である．共感要因の主効果は認められなかったが，交互作用が見つかったため，以下では主効果の説明を省略する．表 4 は 11 項目のアンケートの多重比較の結果を示す．

信頼の結果は共感と成功－失敗系列の間に交互作用があることを示し，多重比較の結果から図 4(a) のよう

表 2: 参加者の信頼に関する情報

成功－失敗系列	共感	平均	標準偏差	成功－失敗系列	共感	平均	標準偏差
	共感あり	56.28	11.03		共感あり	30.53	12.27
フェーズ 1	共感なし	57.70	9.990	フェーズ 4	共感なし	25.19	11.59
	共感あり	26.89	11.57		共感あり	53.77	12.99
フェーズ 2	共感なし	21.89	10.79	フェーズ 5	共感なし	57.88	12.43
	共感あり	57.05	11.52				
フェーズ 3	共感なし	60.95	11.33				

表 3: ANOVA の結果

	要因	F	p	η_p^2
	共感	0.0284	0.8663 ns	0.0001
信頼	成功－失敗系列	608.1	0.0000 ***	0.7563
	共感 × 成功－失敗系列	11.30	0.0000 ***	0.0545

p: *p<0.05 **p<0.01 ***p<0.001

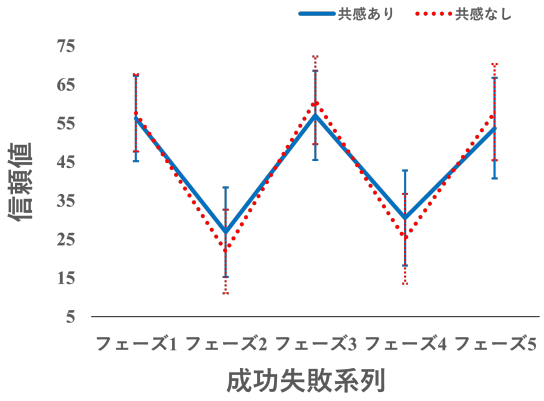

図 3: 共感と成功－失敗系列の交互作用の結果

に，共感を伴う成功-失敗系列因子の単純主効果は 5 段階の組み合わせ間で複数の有意差を示すことが明らかになった．共感を伴わない成功-失敗系列因子の単純主効果も図 4(b) に示すように，5 段階の組み合わせ間で複数の有意を示した．

共感要因のフェーズごとの単純主効果は，基準である第 1 相を除いて，フェーズ 2 からフェーズ 5 まで有意を示した．フェーズ 1 の信頼を基準として用いると，これらの結果はエージェントが共感行動をする場合，そうでない場合よりも経時的な信頼が安定していることを示した．一方，経時的に有意差があるという結果から，フェーズ間の画像認識タスクの成否は信頼値に影響を及ぼさず，エージェントに対する信頼の評価は各フェーズの成否によって異なることが示された．事後分析の結果，共感要因は適切な信頼の構築に有効であることが示された．

5 議論

人間とエージェントの間に信頼関係を適切に構築する方法は，エージェントのパフォーマンスに適した信頼レベルを実現することである．この考えはいくつかの先行研究によって裏付けられている．エージェントへの信頼はエージェントが社会で活用されるために必要な要素であり，適切なアプローチでエージェントへの信頼を一定にすることができれば，人間とエージェントは信頼関係を築くことができる．

本研究では人間がエージェントを信頼するために必要な条件を調査する実験を行った．信頼に影響を与える要因として，エージェントから人間への共感とエージェントの能力の経時的な開示に焦点を当てた．そのため，本研究の目的は共感と成功・失敗の系列要因が，信頼エージェントとのインタラクションによって信頼を制御できるかどうかを調査することである．我々は 2 つの仮説を立て実験から得られたデータを分析した．

実験の結果，共感と成功－失敗系列の交互作用が認められたため，単純主効果を調査するために多重比較行い，共感が存在するフェーズ 1 を基準に，フェーズ 2 からフェーズ 5 にかけて信頼が安定していることが明らかになった．これらの結果は H1 である「エージェントが共感行動をするとき，人間からの信頼は共感行動がない場合よりも安定する」ことを支持した．

さらに、H2 である「信頼の修復はエージェントのミスの後に成功した場合にもとに戻る」ことについても支持された．実験では各フェーズで参加者の信頼度が大きく変化し，直前にミスを犯したエージェントが正解することで，参加者の信頼度が高まった．この実験における直前のタスクの成功または失敗はその時点での信頼であり，これらの信頼は経時的であっても直前の結果を最優先で評価されることが示された．

6 まとめ

エージェントに対する過信や不信の問題を解決するためには，擬人化エージェントと人間の適切な信頼関係の構築が重要な課題である．人間がエージェントとタスクを分担することで，適切な信頼関係がエージェントを人間社会でより活用できるようになることが期

表 4: 単純主効果の結果

	要因	F	p	η_p^2
信頼 (Qt1-11)	フェーズ 1 のときの共感	0.8943	0.3455 *ns*	0.0045
	フェーズ 2 のときの共感	9.891	0.0019 **	0.0480
	フェーズ 3 のときの共感	5.762	0.0173 *	0.0286
	フェーズ 4 のときの共感	9.889	0.0019 **	0.0480
	フェーズ 5 のときの共感	5.176	0.0240 *	0.0257
	共感ありのときの成功 – 失敗系列	235.7	0.0000 ***	0.7063
	共感なしのときの成功 – 失敗系列	379.9	0.0000 ***	0.7949

p: *p<0.05 **p<0.01 ***p<0.001

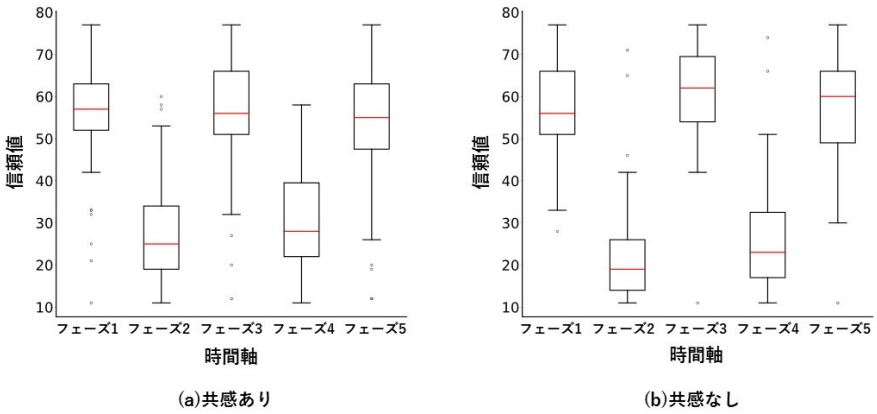

図 4: 信頼尺度の共感要因に関する成功-失敗系列の結果．赤線は中央値，丸印は外れ値．

待できる．本研究はエージェントへの信頼に影響を与える要因の調査の一例であり，実験は参加者間の要因として共感，参加者内要因として成功-失敗系列を持つ 2 要因混合計画で実施した．各要因の水準数は共感 (あり，なし) と成功 – 失敗系列 (フェーズ 1 からフェーズ 5) であった．従属変数はエージェントに対する信頼度であった．結果，共感と成功 – 失敗の系列には交互作用があり，エージェントが共感行動を行うと，フェーズ 1 の信頼値に対して信頼値が修復され，統計的に有意な差があることが示された．これは我々の仮説を支持しており，人間がエージェントを信頼した場合に，共感と成功 – 失敗系列がどのように機能するかを示す重要な例となった．今後の研究として，認知的および感情的な信頼に対する特定の信頼を強めたり弱めたりするケースを検証し，さまざまな状況に対する信頼エージェントの開発を行う．

参考文献

[1] Hannah Bunting, Jennifer Gaskell, and Gerry Stoker. Trust, mistrust and distrust: A gendered perspective on meanings and measurements. *Frontiers in Political Science*, 3, 2021.

[2] Akihiro Maehigashi, Takahiro Tsumura, and Seiji Yamada. Effects of beep-sound timings on trust dynamics in human-robot interaction. In Filippo Cavallo, John-John Cabibihan, Laura Fiorini, Alessandra Sorrentino, Hongsheng He, Xiaorui Liu, Yoshio Matsumoto, and Shuzhi Sam Ge, editors, *Social Robotics*, pages 652–662, Cham, 2022. Springer Nature Switzerland.

[3] Akihiro Maehigashi, Takahiro Tsumura, and Seiji Yamada. Experimental investigation of trust in anthropomorphic agents as task partners. In *Proceedings of the 10th International Conference on Human-Agent Interaction*, HAI '22, page 302–305, New York, NY, USA, 2022. Association for Computing Machinery.

[4] 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.

[5] Tatsuya Nomura, Takayuki Kanda, Hiroyoshi Kidokoro, Yoshitaka Suehiro, and Sachie Yamada. Why do children abuse robots? *Interaction Studies*, 17(3):347–369, 2016.

[6] B. L. Omdahl. *Cognitive appraisal, emotion, and empathy*. Psychology Press, New York, 1 edition, 1995.

[7] Stephanie D. Preston and Frans B. M. de Waal. Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences*, 25(1):1–20, 2002.

[8] Akihiro Maehigashi. The nature of trust in communication robots: Through comparison with trusts in other people and ai systems. In *2022 17th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 900–903, 2022.

[9] Kazuo Okamura and Seiji Yamada. Adaptive trust calibration for human-AI collaboration. *PLOS ONE*, 15(2):1–20, 2020.

[10] Sarah Strohkorb Sebo, Priyanka Krishnamurthi, and Brian Scassellati. " i don't believe you ": Investigating the effects of robot trust violation and repair. In *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 57–65, 2019.

[11] Connor Esterwood and Lionel P. Robert. The theory of mind and human–robot trust repair. *Scientific Reports*, 13(1):9877, Jun 2023.

[12] Xinyi Zhang, Sun Kyong Lee, Whani Kim, and Sowon Hahn. " sorry, it was my fault ": Repairing trust in human-robot interactions. *International Journal of Human-Computer Studies*, 175:103031, 2023.

[13] Takahiro Tsumura and Seiji Yamada. Influence of agent ' s self-disclosure on human empathy. *PLOS ONE*, 18(5):1–24, 05 2023.

[14] Takahiro Tsumura and Seiji Yamada. Influence of anthropomorphic agent on human empathy through games. *IEEE Access*, 11:40412–40429, 2023.

[15] Ana Paiva, Iolanda Leite, Hana Boukricha, and Ipke Wachsmuth. Empathy in virtual agents and robots: A survey. *ACM Trans. Interact. Intell. Syst.*, 7(3), 2017.

[16] Daniel Ullman and Bertram F. Malle. Measuring gains and losses in human-robot trust: Evidence for differentiable components of trust. In *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI), pages 618–619, 2019.

[17] Sherrie Y. X. Komiak and Izak Benbasat. The effects of personalizaion and familiarity on trust and adoption of recommendation agents. *MIS Q.*, 30:941–960, 2006.