

# 既知人物フレーミング：関係性の深い人物を投影させた 対話システムは関係性を引き継ぐ

**Known Individuals Framing: A dialogue system that projects a person with whom one has a close relationship will carry on that relationship.**

角森 穂佳<sup>1\*</sup> 古賀 日南乃<sup>1</sup> 飯田 愛結<sup>1</sup> 奥岡 耕平<sup>1</sup> 佐々木 康輔<sup>2</sup>  
野田 尚志<sup>2</sup> 森口 昌和<sup>2</sup> 大澤 正彦<sup>1</sup>

Honoka Tsunomori<sup>1</sup>, Hinano Koga<sup>1</sup>, Ayu Iida<sup>1</sup>, Kohei Okuoka<sup>1</sup>, Kosuke Sasaki<sup>2</sup>,  
Hisashi Noda<sup>2</sup>, Masakazu Moriguchi<sup>2</sup>, and Masahiko Osawa<sup>1</sup>

<sup>1</sup> 日本大学

<sup>1</sup> Nihon University

<sup>2</sup> NEC ソリューションイノベータ株式会社

<sup>2</sup> NEC Solution Innovators, Ltd.

**Abstract:** 大規模言語モデルの発展により特定の人物を模倣した対話システムの開発が進んでおり、模倣元の人物に対する印象をシステムに投影させることができる。しかし、既知人物を模倣しているという情報提示（フレーミング）自体の影響は明らかではない。そこで、授業支援対話システムを用いて検証した結果、授業担当教員と自身の関係性を高く評価した学生において、専門家よりも授業担当教員としてフレーミングされたシステムへの関係性の印象評価が有意に上昇した。

## 1 はじめに

近年、大規模言語モデルの発展に伴い、ユーザにとって既知である人物（既知人物）を模倣した対話システムの開発が現実的なものとなってきた [24]。既知人物の発話内容を模倣することで、ユーザは対話システムに対して模倣元となった人物に対する印象やその人物との間に存在する関係性を対話システムに投影することが可能になる [4]。

一方で、対話システムに対する印象を変化させる要因は発話内容だけでなく、ユーザがどのような性格や役割をもつ人物像として対話システムを認識させるかも重要な要因となる。こうした人物像をユーザへの情報提示によって印象づける行為はフレーミングと呼ばれ、ユーザの対話システムに対する印象や関わり方を変化させることができる [19, 9]。例えば、ヘルスケアを支援する音声アシスタントにおいて、システムを専門家としてフレーミングするだけで、ユーザからの信頼感を向上させることができる [25]。

しかし、既知人物を模倣した対話システムに関する既存研究では、発話内容の模倣と同時に、既知人物である

というフレーミングも行われることが多く、両者の影響が交絡している。そのため、既知人物としてユーザへ情報提示する、既知人物フレーミングが与える影響については明らかになっていない。また、フレーミングを扱った研究の多くは医者や専門家といった役職のフレーミングの効果を検証しており、既知人物フレーミングが与える影響については、十分に検証されていない。

そこで本研究では、既知人物フレーミングがユーザに与える影響を明らかにするために、探究学習におけるメンタリングを題材とした調査を行った。探究学習におけるメンタリングでは教育者と学習者の間に良好な関係性が築かれていることが求められる。そのため、関係性が既に構築された教育者として学習者にフレーミングすることによって、対話システムに対する関係性と専門性の印象の両方を向上させ、メンタリングに対しても好影響を与えることが期待される。具体的には、大規模言語モデルを用いて探究学習におけるメンタリングを行う対話システムを作成し、実際に授業を受講する履修者を対象にシステムの提供を行った。その際、対話システムを授業の担当教員としてフレーミングする場合と、授業に関連する知識を持った専門家としてフレーミングする場合を比較し、フレーミングがユーザに与える影響を分析した。

\*連絡先：日本大学文理学部

〒156-8550 東京都世田谷区桜上水 3-25-40

E-mail: chho23028@g.nihon-u.ac.jp

## 2 関連研究

### 2.1 特定の人物を模倣する対話システム

特定の人物を模倣する対話システムの既存研究として、様々な模倣の実装方法が提案されている。例えば、Shao らは、有名人のプロフィールや経験、感情状態に基づいて LLaMA-7b モデルを訓練させることで、特定の個人の行動や感情を模倣する、Character-LLM と呼ばれる手法を提案している [21]。また、Tsubota らは大規模言語モデルを用いて特定の人物のプロフィール情報と少数質問で、追加の学習を行わずにプロンプトエンジニアリングのみで高い精度で模倣させる方法を提案している [23]。

模倣を実現する方法ではなく、模倣によってユーザにどのような影響を与えるかに着目した研究として、Hwang らはユーザの家族や主治医といった、既に関係性を構築した実在の人物の人物像をシステムに投影するアプローチを提案している [4]。具体的には、医療カウンセリングの文脈において、家族と主治医という既知人物の情報を基に、アカウントのアイコンや語り口を似せたチャットボットを作成してユーザに与える効果を検証している。調査の結果、家族を模したシステムがユーザの受容度を高め、主治医を模したシステムでは信頼感を向上させることを明らかにした。しかし、これらの既存研究では対話内容の模倣による効果とフレーミングを行うことの効果とが混在しており、フレーミングによる効果が明らかになっていない。

### 2.2 対話システムにおけるフレーミング

対話システムにおけるフレーミングを用いた研究で、盛んに扱われているのが専門性のフレーミングである [25]。例えば、Liew らは会話型コマースの文脈において、専門性を想起させるフレーミングを行うとチャットボットの信頼や購買意図が高まることを報告している [13]。

フレーミングする人物像は専門家像以外にも広範にわたる [12]。例えば、Rhee らは商品推薦の場面において音声エージェントを友人役としてフレーミングすると、製品推奨の受容性が高まると報告した [18]。また、Baylor らは、教育支援を目的とした対話システムにおいて、専門家、学習者を励ますモチベーター、励ましと指導の両方を行うメンターの3種類のフレーミングを比較し、異なる役割提示が自己効力感やシステムに対する信頼感に影響することを示している [1]。

しかし、既存のフレーミング研究では専門家や友人といった、役割を提示するフレーミングを主に扱っており、既知人物としてフレーミングすることによる影響は調査されていない。既知人物としてフレーミングする

ことで、既存の単なる役割提示に比べて、ラポールや信頼関係といった長期的な相互作用を通じて構築される印象を向上させる可能性がある。

### 2.3 探究学習におけるメンタリング

探究学習とは、学習者自らが問題を設定し、設定した問題を解決するプロセスを経験することで、課題解決能力などを習得する学習法である。近年、探究学習に対する関心が高まっており、大学等の高等教育機関でもその実践が求められている [22, 11]。探究学習において教育者は、学習者に対する質問などの対話を通して学習者の関心や問題に対する理解を引き出し、思考を促すことが求められる [6]。

このような教育における教育者と学習者の関係性はメンターシップとも呼ばれ [5]、教育者からの思考を促すことを目的とした対話はメンタリングの1種と位置付けられる。効果的なメンタリングを実現するためには、メンター（教育者）とメンティ（学習者）の間の良好な関係性が築かれていることが重要とされている [20]。

本研究では、既知人物フレーミングの効果の検証をするために、探究学習におけるメンタリングを題材として調査を行った。関係性が既に構築された教育者として学習者にフレーミングすることによって、対話エージェントに対する関係性と専門性の印象の両方を向上させ、効果的なメンタリングを実現することができるかを検証する。

## 3 授業での調査と分析

### 3.1 目的

本調査の目的は、以下の2つの問いを明らかにすることである。

**RQ1** 既知人物フレーミングのみによって、対話システムに既知人物との関係性を投影させることが可能か

**RQ2** 既知人物フレーミングによる関係性の投影が、メンタリングに影響を与えるか

本調査は日本大学文理学部研究倫理委員会の承認を受けて実施した（承認番号:07-76）。なお、研究と教育の分離について十分に考慮し、本システムは授業の学習効果を高めるための教材として全履修者を対象に導入した。一方で、研究への同意の有無が成績評価や学習機会に影響しないことを事前に説明し、同意が得られた履修者のデータのみを授業後に抽出し分析対象とした。

### 3.2 設定

システムの提供は、探究学習のメンタリングを日本大学文理学部内で開講されている授業の履修者を対象として授業内で行った。授業では、OKR(Objectives and Key Results) と呼ばれる目標を設定・管理するためのフレームワークを用いていた [3]。OKR は、達成すべき目標「Objective」とその目標の達成度を測るための定量的で測定可能な指標である「Key Results」で構成されている。授業では、履修者がより良い OKR を設定できるように自らの OKR に対する思考を促すことを目的に、システムとのメンタリングを行ってもらった。

システムは、授業内で用いられていたチャットツールである Slack 上で稼働する Slack ボットとして実装した。システムは大規模言語モデルを用いて履修者の発話に対して回答を生成し、返答を行う。なお、本調査ではフレーミングによる効果に着目するため、すべての条件で同様のシステムを用いて対話を行った。大規模言語モデルに与えたプロンプトの概要を以下に示す。

#### # 概要

あなたは大学の Slack ワークスペースで動作するメンタリング用チャットボットです。学生達はそれぞれ Objective and Key Results (OKR) という目標設定フレームワークに基づいて自らの目標を設定しているので、より良い OKR を設定できるように、下記の説明に基づいてアドバイスをしてください。

#### # Objective and Key Results について (OKR に関する説明)

#### # 制約条件

(話し方に対する制約)

(アドバイスの戦略に関する制約)

概要ではシステムの役割について概説を行い、OKR に関する説明の部分では、OKR 自身の説明やどのような OKR が良いかといった説明を記載した。なお、OKR に関する説明は授業資料から抜粋して作成した。最後に、制約条件の部分では発話量や口調といった話し方や、アドバイスしすぎずに学生からの意見を引き出すように話すといったメンタリングにおけるアドバイスの仕方について制約を記載した。なお、本調査では大規模言語モデルとして OpenAI の gpt-4.1-mini[17] を用いた。

### 3.3 条件

本調査は、学生にとって既知の授業を担当する教員としてフレーミングする既知人物フレーミング条件と、

#### メンタリングAIからの自己紹介



こんにちは、私は大澤先生AIです。大澤先生の知見を踏まえて、大澤先生の代わりに皆さんにメンタリングを行います。よろしくお願いします。

(a) 既知人物フレーミング条件

#### メンタリングAIからの自己紹介



こんにちは、私はOKR専門家AIです。OKR専門家の知見を踏まえて、OKR専門家の代わりに皆さんにメンタリングを行います。よろしくお願いします。

(b) 専門家フレーミング条件

図 1: 各条件下の教示内容 (名称・アイコン・自己紹介)

一般的な専門家としてフレーミングする専門家フレーミング条件の 2 つの条件を用意した。履修者に各条件で教示する内容は、既存研究 [4, 12, 25] を参考に、システムの名称、アイコン、自己紹介、説明文の 4 つの項目について教示を設計した。

図 1 に示すように、名称は既知人物フレーミング条件ではシステムの名称を「(担当教員の苗字) 先生 AI」に、専門家フレーミング条件では「OKR 専門家 AI」として提示した。また、アイコンは既知人物フレーミング条件では担当教員の画像を、専門家フレーミング条件では博士帽を被った人物を抽象化したアイコンを用いた。自己紹介は、「こんにちは、私は【システム名】。【対象人物像】の知見を踏まえて、【対象人物像】の代わりに皆さんにメンタリングを行います。よろしくお願いします。」といった定型文を用意し、【システム名】の部分にはシステムの名称を、【対象人物像】の部分には既知人物フレーミング条件では「(担当教員の苗字) 先生」、専門家フレーミング条件では「OKR 専門家」と置き換えて提示した。

図 1 に示す 3 項目に加えて、システムの説明文として「今回のメンタリングで使用するシステムは、OKR を専門に研究する【対象人物像】の知見や【対象人物像】の対話データから学習して開発した AI です。今回のメンタリングでは、この AI が【対象人物像】の代わりにメンタリングを行います。」といった定型文を用意し、同様に【対象人物像】の部分に条件に沿って置き換えて提示した。

なお、システムの提供およびその体験は大学の授業として実施しており、全履修者が両方のシステムを体験する授業設計であったため、教育機会の均等は担保されている。ただし、体験の順序は履修者ごとにランダム

に割り当て、研究利用に同意した履修者のうち 1 回目に体験したデータのみを抽出して分析対象とした。

### 3.4 評価方法

本調査では既知人物フレーミングによる影響を明らかにするため、システムに対する関係性と専門性の印象およびメンタリング効果に関するアンケート評価を行った。

関係性の評価には、メンタリングにおける関係性を評価する尺度であるメンターシップ・ワーキング・アライアンス (MWA) 尺度の一部を和訳して用いた [20]。MWA 尺度の質問はメンティがメンターに対する絆や信頼関係といった関係性の質を評価する項目と、メンターの行動を評価する項目の大きく 2 つの項目で構成されている。本研究では、フレーミングによってシステムに対してどのような関係性を認識しているか評価するために、関係性の質に該当する質問項目を用いた。質問項目は 5 項目で、各項目を 5 段階のリッカート尺度で回答してもらった値を平均化した値を関係性に対する印象を示す指数とした。加えて独自項目として、システムに対して担当教員らしさを感じたか、システムへの回答には担当教員の知見が反映されていると感じたかを 5 段階のリッカート尺度で回答してもらった。なお本調査は、関係性の評価値が参加者の平均以上であった参加者に対して分析を行った。これは、事前に担当教員との関係性が構築されていない場合、既知人物フレーミングによる影響が現れないと考えられるためである。

システムに対する専門性の印象評価には、McCroskey らの SD 法に基づく質問紙を用いた [16]。質問紙は 6 項目の形容詞対からなり、各項目を 7 段階で回答してもらった値を平均化し、専門性の印象を示す指数とした。

次に、メンタリングにおける既知人物フレーミングの有用性を評価するため、メンタリング自体に対する印象とメンタリングの目的である OKR に対する影響の 2 つの観点から評価した。メンタリング自体に対する印象については、メンターからの支援の度合いを評価する日本語版 Mentoring Functions Questionnaire (MFQ-9) を用いた [28]。また、OKR に対する影響の評価では、メンタリングの前後で参加者の OKR を回答してもらい、メンタリング後には OKR を変更したかを尋ねることで OKR の変更の有無を調査した。加えて、設定した OKR に対する印象の変化を評価するために、目標設定理論 [14] に基づく既存尺度を参考に目標の明確性 [10]、困難性 [10]、およびコミットメント [7] の 3 観点から評価を行った。

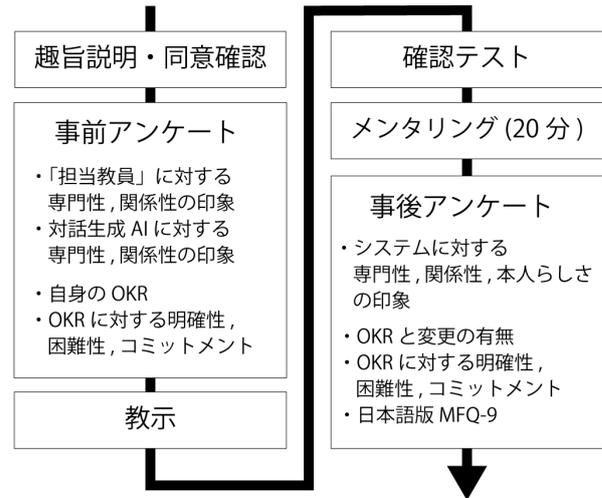


図 2: システム提供を行った授業の流れ

### 3.5 手順

システムの提供を行った授業の流れを図 2 に示す。まず初めに、体験内容とシステムの説明を行いデータの研究利用への同意の確認を行った。その後、履修者の各自の PC でシステムにアクセスしてもらい、事前アンケートに回答してもらった。事前アンケートではまず、担当教員と一般的な生成 AI に対する関係性と専門性の印象を尋ねた。加えて、参加者の OKR を回答してもらい、目標設定である明確性・困難性・コミットメントの 3 つの観点で自身の OKR について回答してもらった。

事前アンケート回答後、各条件に応じたシステムに関する教示を行った。教示後、正しく教示がされているかを確認するテストを行った。テストは選択式の問題で構成されており、両条件共通の質問として、「システムは誰の知見を基にしていると説明されたか」「システムは誰の代わりになってメンタリングを行うと説明されたか」の 2 問を設定した。加えて既知人物フレーミング条件では、「担当教員が何の研究を専門としていると説明されたか」という質問を追加した。これは、担当教員の主たる専門領域は OKR 研究では無かったため、教示文を読まずに主たる領域の専門家としてのみ理解していた参加者を除外するためである。

テスト回答後、授業運営に協力したシステム開発者の合図で一斉にメンタリングを行ってもらった。なお、最初の発話は自らの OKR を含む全参加者共通のテキストとした。メンタリングは 20 分間行ってもらい、20 分が経過した時点で会話を中断し、事後アンケートに回答してもらった。事後アンケートでは、対話したシステムに対する関係性・専門性の印象と、本人らしさに関する 2 つの質問に回答してもらった。加えて、メンタリングに関する評価として、事前アンケートと同様に自

表 1: 事前アンケートにおける担当教員と対話生成 AI に対する関係性と専門性の印象評価

項目	既知人物 M(SD)	専門家 M(SD)
【担当教員に対する印象】		
関係性	4.18 (0.41)	3.95 (0.47)
専門性	6.72 (0.53)	6.53 (0.75)
【対話生成 AI に対する印象】		
関係性	3.51 (0.72)	2.74 (0.82)
専門性	5.51 (1.10)	5.30 (1.17)

表 2: 事後アンケートにおけるシステムに対する印象評価および日本語版 MFQ-9 の結果

項目	既知人物 M(SD)	専門家 M(SD)
【システムに対する印象】		
関係性	3.99 (0.43)	3.45 (0.78)
専門性	5.99 (0.60)	5.56 (1.01)
本人らしさ	4.50 (0.52)	3.39 (1.04)
知見の反映	4.50 (0.90)	3.67 (0.77)
【日本語版 MFQ-9】		
キャリア支援	4.19 (0.61)	3.87 (0.85)
心理社会的支援	2.53 (0.89)	2.28 (1.16)
ロールモデル支援	3.31 (0.48)	3.26 (1.05)

身の OKR と明確性・困難性・コミットメントの評価、また OKR の変更の有無と日本語版 MFQ-9 に回答してもらった。

### 3.6 結果

参加者 65 名の内、Manipulation Check としてメンタリング直前に行ったテストに全て正答し、担当教員に対する関係性の評価値が参加者の平均の 3.42 以上であった参加者を分析対象者とした。分析対象者は既知人物フレーミング条件 12 名、専門家フレーミング条件 18 名の計 30 名であった。事前アンケートの結果を表 1 に、事後アンケートの結果を表 2 に示す。

まず、フレーミングによって参加者がシステムに対して担当教員を想起していたか確かめるために、事後アンケートにおける本人らしさの評価項目についてそれぞれウィルコクソンの順位和検定を行った所、本人らしさ ( $p = .005, r = -0.51$ ) および知見の反映 ( $p = .008, r = -0.49$ ) の両方について有意差が見られた。このことから、既知人物フレーミングによってシステムに対して専門家フレーミングに比べて担当教員を想起させていた

表 3: 目標設定 (明確性・困難性・コミットメント) に関する評価

項目	既知人物 M(SD)	専門家 M(SD)
【事前】		
明確性	3.83 (0.80)	4.03 (0.70)
困難性	4.67 (0.29)	4.46 (0.55)
コミットメント	3.20 (0.87)	3.61 (0.88)
【事後】		
明確性	3.96 (0.85)	4.19 (0.82)
困難性	4.44 (0.37)	4.47 (0.57)
コミットメント	3.50 (0.66)	3.66 (0.75)

表 4: 各条件下の OKR を変更した参加者の割合

項目	既知人物	専門家
Objective	33.33%	38.89%
Key Results	66.67%	38.89%

ことが示された。

次に、事後アンケートにおける関係性について条件間でウィルコクソンの順位和検定を行った結果、関係性について既知人物フレーミング条件が有意に関係性に対する印象が深いことが示された ( $p = 0.04, r = -0.38$ )。

また、専門性については、専門家フレーミングと同等の専門性を想起させたか検証するために、同等性検定としてウィルコクソンの順位和検定を用いた Two One-Sided Tests (TOST) を、同等性マージンを  $\delta = 0.5$  として行った。結果、下側境界 ( $\theta > -\delta$ ) については有意であったが ( $p < 0.01$ )、上側境界 ( $\theta < +\delta$ ) については有意に棄却できなかった ( $p = 0.46$ )。そのため、同等性マージンを  $\delta = 0.5$  内での同等性は支持されなかった。

加えて、メンタリングに関する評価項目として、日本語版 MFQ-9 と目標設定に関する項目の結果を分析した。まず、日本語版 MFQ-9 の結果について 3 つの下位尺度についてそれぞれウィルコクソンの順位和検定を行った所、キャリア的支援 ( $p = 0.34, r = -0.17$ )、心理社会的支援 ( $p = 0.42, r = -0.15$ )、ロールモデル的支援 ( $p = 1.00, r = 0.00$ ) のいずれの尺度についても有意差は見られなかった。

続いて、目標設定に関する 3 項目の結果を表 3 に示す。3 つの項目について、条件間の違いを水準とする「フレーミング要因」と、メンタリングの前後を水準とする「メンタリング経験要因」の 2 要因において、整列ランク変換を施した二元配置分散分析 [26] を実施したところ、いずれの項目にも交互作用は見られなかった (明確性: $p = 0.699, \eta_p^2 < .01$ ; 困難性: $p = 0.105, \eta_p^2 =$

0.09; コミットメント: $p = 0.125, \eta_p^2 = 0.08$ ). そこで各要因の主効果を見ると、フレーミング要因については有意差は見られなかった(明確性: $p = 0.502, \eta_p^2 = 0.02$ ; 困難性: $p = 0.908, \eta_p^2 < 0.01$ ; コミットメント: $p = 0.320, \eta_p^2 = 0.04$ ). メンタリング経験要因については、困難性とコミットメントについて有意差は見られなかったものの(困難性: $p = 0.267, \eta_p^2 = 0.04$ ; コミットメント: $p = 0.213, \eta_p^2 = 0.05$ ); 明確性については有意にメンタリング後に向上していた( $p = 0.046, \eta_p^2 = 0.13$ ). このことから、両条件においてメンタリングによって明確性が向上したことが示された。

最後に、各条件における Objective と Key Results を変更した参加者の割合を表4に示す。フィッシャーの正確確率検定を用いて条件間で比較した所、Objective( $p = 1.00, \phi = 0.06$ )と Key Results( $p = 0.26, \phi = 0.27$ )の両方で有意差は見られなかった。

## 4 考察

調査の結果、既知人物フレーミングによって対話システムに対する関係性の印象が有意に向上した。このことから、RQ1として掲げた「既知人物フレーミングのみによって対話システムに既知人物との関係性を投影させることが可能か」という問いに対して、その可能性が示唆された。

一方で、専門性の評価については条件間で同程度の専門性知覚の評価値は見られなかった。結果としては既知人物フレーミング条件の方が専門性の平均値は高く、上側境界 ( $\theta < +\delta$ ) について有意に棄却できなかったことから、既知人物フレーミングが専門家フレーミングに比べて専門性をやや高く評価される傾向にあった可能性がある。原因の一つとして、担当教員の方が授業資料など専門性の根拠となる情報に触れる機会が多く、専門性が高く評価されていた可能性がある。Bergmanらは、専門性の根拠となる情報提示が多いほど専門性の知覚を向上させることを示しており [2]、本結果はこの既存研究の知見と整合的である。また、Yangらは親しみやすいエージェントほど知識能力が高く認識される傾向があることを示しており [27]、投影された関係性の向上によって親しみやすさが向上したことも、専門性に対する知覚が向上した一因として考えられる。

また、メンタリングの効果に関しては、メンタリングの前後で明確性のみが有意に向上し、フレーミング要因間では有意な差は認められなかった。このことから、RQ2として設定した「既知人物フレーミングによる関係性の投影が、メンタリングに影響を与えるか」という問いに対して、少なくとも本調査の範囲では、影響を与えたとは言えない結果となった。メンタリングの前後で明確性が向上した要因としては、エージェントに対

して目標に関する思考を言語化する過程が、目標の明確化を促進したためであると考えられる。この結果は、エージェントとの対話によって新たな気づきを促した事例 [8] や、目標の具体性を向上させた事例 [15] といった既存研究の知見と一致する。

## 5 おわりに

本研究では、対話システムにおける既知人物フレーミングの効果を検証した。探究学習におけるメンタリングを行う対話システムを用いて調査した結果、専門家という役割でフレーミングする場合に比べ、担当教員という関係性を築いた既知人物としてフレーミングすることで、ユーザは既知人物との関係性をシステムに投影できることが明らかとなった。

本研究の知見は、特定の個人を再現した対話システムのデザインにおいて2つの重要な示唆をもたらす。第一に、実装コストの低減である。高精度な模倣を実現せずとも、適切な既知人物フレーミングによって関係性の投影が実現できるため、個人を再現した対話システム作成のハードルを大きく下げる可能性がある。第二に、実社会応用におけるリスク対策である。本調査では、システムがAIであることを明記しているにも関わらず、模倣元の人物との関係性を投影していた。このことから、エージェントの不適切な応答がモデルとなった本人への直接的な失望に繋がるリスクを回避しつつ、親近感や信頼感といった関係性のメリットのみを享受できるインタラクションデザインを実現できる可能性がある。

## 参考文献

- [1] Amy L. Baylor and Yanghee Kim. Simulating instructional roles through pedagogical agents. *International Journal of Artificial Intelligence in Education*, Vol. 15, No. 2, pp. 95–115, 2005.
- [2] Abigail Bergman, Mohamed A. Hussein, Rhia Catapano, and Zakary L. Tormala. Fifteen reasons you should read this paper: How providing many arguments increases perceptions of both expertise and persuasive intent. *Personality and Social Psychology Bulletin*, p. 01461672251366068, 2025.
- [3] J. Doerr and L. Page. *Measure What Matters: How Google, Bono, and the Gates Foundation Rock the World with OKRs*. 2018.
- [4] Youjin Hwang, Donghoon Shin, Sion Baek, Bongwon Suh, and Joonhwan Lee. Applying the

- persona of user's family member and the doctor to the conversational agents for healthcare. *CoRR*, Vol. abs/2109.01729, , 2021.
- [5] W. Brad Johnson. The intentional mentor: Strategies and guidelines for the practice of mentoring. *Professional Psychology: Research and Practice*, Vol. 33, No. 1, pp. 88–96, 2002.
- [6] Aisha Kawalkar and Jyotsna Vijapurkar. Scaffolding science talk: The role of teachers' questions in the inquiry classroom. *International Journal of Science Education*, Vol. 35, No. 12, pp. 2004–2027, 2013.
- [7] Howard J. Klein, Michael J. Wesson, John R. Hollenbeck, Patrick M. Wright, and Richard P. DeShon. The assessment of goal commitment: A measurement model meta-analysis. *Organizational Behavior and Human Decision Processes*, Vol. 85, No. 1, pp. 32–55, 2001.
- [8] Rafal Kocielnik, Lilian Xiao, Daniel Avrahami, and Gary Hsieh. Reflection companion: A conversational system for engaging users in reflection on physical activity. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol*, Vol. 2, No. 2, pp. 1–26, 2018.
- [9] Philipp Krop, Martin Jakobus Koch, Astrid Carolus, Marc Erich Latoschik, and Carolin Wienrich. The effects of expertise, humanness, and congruence on perceived trust, warmth, competence and intention to use embodied ai. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, pp. 1–9, 2024.
- [10] Ho Kwong Kwan, Cynthia Lee, Phyllis L. Wright, and Chun Hui. Re-examining the goal-setting questionnaire. In *New Developments in Goal Setting and Task Performance*. 2013.
- [11] Philippa Levy and Robert Petrulis. How do first-year university students experience inquiry and research, and what are the implications for the practice of inquiry-based learning? *Studies in Higher Education*, Vol. 37, No. 1, pp. 85–101, 2012.
- [12] Tze Wei Liew, Su-Mae Tan, and Chandrika Jayothisa. The effects of peer-like and expert-like pedagogical agents on learners' agent perceptions, task-related attitudes, and learning achievement. *Journal of Educational Technology & Society*, Vol. 16, No. 4, pp. 275–286, 2013.
- [13] Tze Wei Liew, Su-Mae Tan, Jessica Tee, and Gerald Guan Gan Goh. The effects of designing conversational commerce chatbots with expertise cues. In *2021 14th International Conference on Human System Interaction (HSI)*, pp. 1–6, 2021.
- [14] Edwin A. Locke. Toward a theory of task motivation and incentives. *Organizational Behavior and Human Performance*, Vol. 3, No. 2, pp. 157–189, 1968.
- [15] Gabrielle Martins Van Jaarsveld, Jacqueline Wong, Martine Baars, Marcus Specht, and Fred Paas. Scaling goal-setting interventions in higher education using a conversational agent: Examining the effectiveness of guidance and adaptive feedback. In *Proceedings of the 15th International Learning Analytics and Knowledge Conference*, pp. 328–338, 2025.
- [16] James C. McCroskey and Jason J. Teven. Goodwill: A reexamination of the construct and its measurement. *Communication Monographs*, Vol. 66, No. 1, pp. 90–103, 1999.
- [17] OpenAI. Gpt-4.1-mini, 2024.
- [18] Chong Eun Rhee and Junho Choi. Effects of personalization and social role in voice shopping: An experimental study on product recommendation by a conversational voice agent. *Computers in Human Behavior*, Vol. 109, p. 106359, 2020.
- [19] Minjin Rheu, Yue Dai, Jingbo Meng, and Wei Peng. When a chatbot disappoints you: Expectancy violation in human-chatbot interaction in a social support context. *Communication Research*, Vol. 51, No. 7, pp. 782–814, 2024.
- [20] Jenna Griebel Rogers and Angela Byars-Winston. Measuring the interpersonal component of the mentoring relationship: The mentorship working alliance scale – mentee version. *Journal of Clinical and Translational Science*, Vol. 9, No. 1, p. e99, 2025.
- [21] Yunfan Shao, Linyang Li, Junqi Dai, and Xipeng Qiu. Character-llm: A trainable agent for role-playing. In *EMNLP*, 2023.

- [22] Melissah B. Thomas, Amanda Muscat, Ashleigh Zuccolo, Carla Nascimento Luguetti, and Anthony Watt. Navigating pedagogical innovation in higher education: Education academics' experiences with active and inquiry-based learning in intensive teaching. *Innovative Higher Education*, Vol. 50, No. 6, pp. 1917–1943, 2025.
- [23] Yuka Tsubota and Yoshinobu Kano. Text generation indistinguishable from target person by prompting few examples using llm. In *Proceedings of the 2nd International AIWolfDial Workshop*, pp. 13–20, 2024.
- [24] Noah Wang, Z. Y. Peng, Haoran Que, Jiaheng Liu, Wangchunshu Zhou, Yuhan Wu, Hongcheng Guo, Ruitong Gan, Zehao Ni, Jian Yang, Man Zhang, Zhaoxiang Zhang, Wanli Ouyang, Ke Xu, Wenhao Huang, Jie Fu, and Junran Peng. Rolellm: Benchmarking, eliciting, and enhancing role-playing abilities of large language models. In *Findings of the Association for Computational Linguistics: ACL 2024*, pp. 14743–14777, 2024.
- [25] Carolin Wienrich, Clemens Reitelbach, and Astrid Carolus. The trustworthiness of voice assistants in the context of healthcare. *Frontiers in Computer Science*, Vol. 3, p. 685250, 2021.
- [26] Jacob O. Wobbrock, Leah Findlater, Darren Gergle, and James J. Higgins. The aligned rank transform for nonparametric factorial analyses using only anova procedures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 143–146, 2011.
- [27] Fu-Chia Yang, Siqu Guo, and Christos Mousas. Exploring familiarity and knowledgeability in conversational virtual agents. *ACM Transactions on Applied Perception*, Vol. 23, No. 1, 2025.
- [28] 榑原（関）圭子, 石川ひろの, 木内貴弘. 日本語版 mentoring functions questionnaire 9 項目版 (mfq-9) の信頼性・妥当性の検討. *産業衛生学雑誌*, Vol. 55, No. 4, pp. 125–134, 2013.