



没入型バーチャル環境における講師エージェントの数が記憶に与える影響の予備的調査

瑞穂嵩人¹⁾, 鳴海拓志¹⁾, 葛岡英明¹⁾

1) 東京大学 大学院情報理工学系研究科 (〒 113-0033 東京都文京区本郷 7-3-1, {takato, narumi, kuzuoka}@cyber.t.u-tokyo.ac.jp)

概要: 遠隔講義や非同期教育の効果を高めるため, 講師役のバーチャルエージェントの設計が記憶に与える影響が調査されてきた. 先行研究が主に外見の効果を調査してきたのに対して, 本研究では没入型バーチャル環境内の講師エージェントの数が記憶に与える影響を調査した. 2 人の講師エージェントによる情報提示が, 1 人のエージェントを用いる場合よりも記憶課題の成績を向上するという仮説を立て, 予備的調査を行った. 8 名を対象とした調査の結果, 条件の実施順序で効果が逆転する可能性などが示唆され, 本調査に向けて実験デザインの修正などが必要であると示唆された.

キーワード: バーチャルエージェント, 記憶, 学習, VR

1. はじめに

遠隔講義やオンライン学習の効果を高める手法として, 画面上に講師役のバーチャルエージェントやアバタを表示するアプローチが提案されている. これらのバーチャルエージェントやアバタは, 単なるモニタ上の画像であることを超えて, ユーザとの間に社会的関係性を構築し, ユーザの知覚や認知に様々な影響を与えることが知られている. 例えば, Baylor and Kim [1] は, バーチャルエージェントの人種, 性別, 外見のリアルさ (リアル調・アニメ調), 役割 (専門家・同僚) といった属性が学習に与える影響を検証し, エージェントがリアルなほど学習効果が向上することを示唆した. Mizuho et al. [2] は, ヒト・動物・ロボット・エイリアンの外見を持つキャラクタが話す単語を覚える記憶課題において, ヒトエージェントが他の表現に比べて記憶成績を向上させる可能性を示唆した. このように, エージェントの表現を操作することで, ユーザとの社会的関係性を変化させ, 学習や記憶の支援を試みる研究が行われてきた.

従来, 教育・学習用途のエージェント研究は, 主に 2D モニタ上のエージェントを対象としてきた. 他方, 2010 年代後半以降, 廉価なヘッドマウントディスプレイ (HMD) の普及に伴い, バーチャルリアリティ (VR) を用いた没入型バーチャル環境における講師エージェントの活用が目立ち始めていた. 没入度の向上によりエージェントの社会的存在感が向上することが知られており [3], 2D 環境と比較して VR 環境では, 講師エージェントが学習・記憶に与える影響が強くなると期待されている. 本領域は萌芽の段階にあり, 知る限りでは, 先行研究はまだ数例であるが, 主にエージェントの外見に関する効果が調査されている. 例えば, 外見のリアルさ [4], ユーザの性別に応じた外見 (例: 女子学生には若い女性研究者を模したエージェント) [5], 反復学習における外見の多様化 (例: 医者, 軍人, ビジネス

パーソン) [6] といった操作の効果が検証されている.

このように, VR における講師エージェント研究は, 2D エージェント研究の延長として, 外見の効果に焦点を当ててきた. しかし, VR では 2D 環境と異なり, エージェントの空間的配置や身体動作など, 外見以外の要因も操作可能となる. したがって, VR ならではの手法や効果を探ることが, 2D エージェント研究の単なる拡張にとどまらない新たな知見の獲得につながると考えた. 本研究では, 特にソーシャルロボット研究においてエージェントの効果を高める要因として知られる, エージェントの数に着目する. 先行研究では, 1 台のロボットよりも 2 台のロボットを用いた条件において, ロボットエージェントによる学習の動機づけ [7] や運動スキルの定着効果 [8] が向上することが示唆されている. これらの知見を踏まえ, 本研究では VR において 2 人の講師エージェントを用いる条件が 1 人の講師エージェントを用いるよりも記憶課題の成績が高くなると仮説を立てた. 仮説検証のため, VR 環境下で講師エージェントが外国語-日本語ペアを記憶ターゲットとして提示する実験システムを構築し, 8 名を対象に予備的調査を実施した. 予備的調査の目的は, 記憶課題の難易度の調整 (床効果・天井効果の回避), システム全体の動作確認, および効果量の見積もりなどであった. また, 参加者に対してインタビューも実施し, 2 人の講師エージェントの設計に関する質的知見の収集も試みた.

2. 実験

2.1 実験デザインおよび仮説

参加者は, VR 環境内で講師役のバーチャルエージェントが声で提示するタグログ語-日本語ペアを覚える記憶課題に取り組んだ. 実験は 2 水準参加者内計画で実施され, バーチャルエージェントが 1 体の Single 条件と 2 体の Double 条件が比較された (図 1). 評価指標は, 想起テストの正答

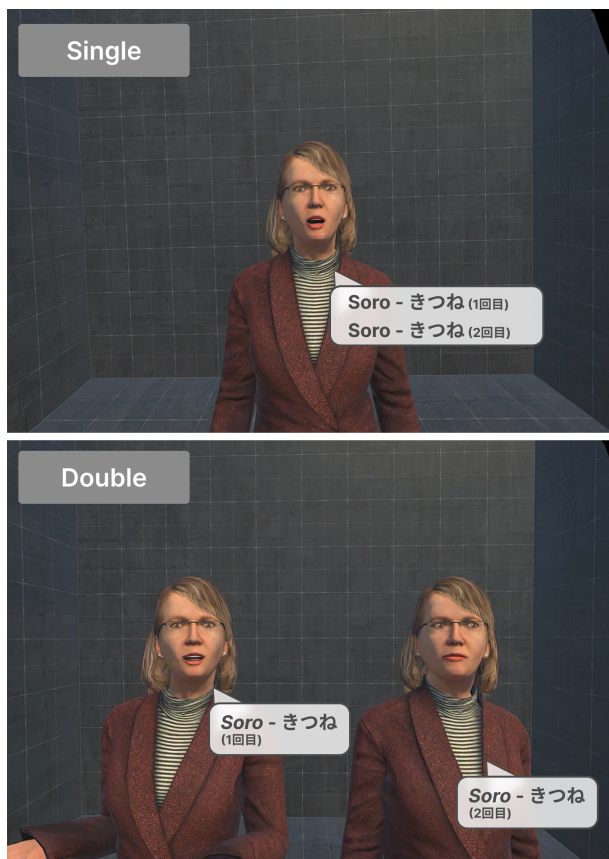


図 1: 実験システムおよび条件. Single 条件では, 1 人の講師エージェントが単語ペアを 2 度復唱した. 他方, Double 条件では 2 人の講師エージェントが分担して, 1 回ずつ単語ペアを発話した.

率, 質問紙で計測されるエージェントの社会的存在感や不気味さであった. 両条件で使用されたエージェントの外見と声は同一であった. また, 教示や単語などのエージェントから提示される情報量は条件間で等価であった. 具体的には, Double 条件では Single 条件の対話スクリプトを 2 体のエージェントが分担して交互に発話した. 条件の実施順序は, 参加者間でカウンタバランスがとられた.

先行研究では, エージェントの数がユーザへの社会的影響を促進することが示唆されている [7, 8]. したがって, 記憶課題の正答率について, 以下のように仮説を立てた.

H1 記憶課題の正答率は, Single 条件よりも Double 条件のほうが高い

また, 同じ外見の人物を複数人同時に見ることで不気味さを感じる効果 (クローン減価効果) が知られている [9]. したがって, エージェントに対して知覚される不気味さについて, 以下のように仮説を立てた.

H2 不気味さの評定は, Single 条件よりも Double 条件のほうが高い

2.2 実験参加者

健常な男女 8 名 (平均 22.6 ± 1.30 (SD) 歳, 男性 7 名, 女性 1 名) が参加した. 記憶ターゲットの一部である日本

語名詞への慣れが記憶成績に影響する可能性を考慮し, 参加者は全員日本語母語話者とした. 参加者は, 実験参加に対する謝礼として, Amazon ギフト券 1500 円分を受け取った.

2.3 記憶ターゲット

参加者は 10 組のタガログ語-日本語ペアを 2 セット (List A, B) 覚えた. タガログ語を採用した理由は, 先行研究 [6] と同様に, アルファベットで記述されるため英語と近い感覚で学習できる点と, ほとんどの参加者にとって事前知識がないという点の実験に適していたためである. 単語間の覚えやすさの違いに寄与する可能性がある単語の長さを統制するため, タガログ語の音節数が 2 の名詞を辞書から 20 語選定した. 例えば, ahas (蛇), bilog (丸) などが含まれていた. どちらの条件でどちらの単語セットを用いるかについては, 参加者間でカウンタバランスがとられた.

2.4 システム

VR 環境は Unity 2021.3 を使用して描画された. 参加者は, HMD (Meta Quest 3) とヘッドホン (Sony WH1000XM3) を装着し, VR 環境に没入した. バーチャルエージェントへの自然な注意を促進するため, VR 環境は家具のないシンプルなバーチャルルームを使用した. VR 環境内で参加者の身体はセルフアバタとして表示された. セルフアバタは, オープンソースである Microsoft Rocketbox Avatar Library [10] から, 特別な意味を持たないニュートラルなアバタとして, 男性参加者用の Male_Adult_08 と女性参加者用の Female_Adult_01 を採用し, 性別に合わせて使用した. HMD と両手の位置・姿勢に基づき, RootMotion FinalIK プラグイン¹を用いた逆運動学計算によりセルフアバタの全身運動が実現された.

講師エージェントも Microsoft Rocketbox Avatar Library から選定された. 先行研究 [6] で使用されたアバタのうち, 参加者からのフィードバックで講師役としてもっともらしいとポジティブに評価されることが多かった Business_Female_02 を採用した. エージェントの身体動作は Microsoft Rocketbox Avatar Library に収録されているアニメーションの中から, 状況に適したもの (例えば, アイドリング動作など) が選定され, 反映された. エージェントの音声は, タガログ語と日本語ともに Ondoku3² で作成され, エージェントの口は OVRLipsync プラグイン³によって音声と同期して動いた. 各バーチャルエージェントの口の位置から声が聞こえるようにバーチャルスピーカを配置し, 空間音響が設定された. これらの実装によって, 特に Double 条件において, どちらのエージェントが喋っているのかを自然に知覚できるように工夫した.

2.5 手順

参加者は, まず実験説明書を読み, 実験参加に同意する旨をオンラインフォームを通じて回答した. 実験手順に関

¹<https://assetstore.unity.com/packages/tools/animation/final-ik-14290>

²<https://ondoku3.com/>

³<https://developer.oculus.com/downloads/package/oculus-lipsync-unity/>

する詳細な説明を受けたあと、VR 酔いを評価するための Simulator Sickness Questionnaire (SSQ) に回答した。その後、HMD とヘッドホンを装着し、VR 環境に没入した。講師エージェントから簡単な挨拶（例：「こんにちは、〇〇さん。」）があり、その後 30 秒間周囲を見回すように指示された。この 30 秒間は VR 環境やエージェントに慣れるための時間であった。続いて、講師エージェントが 10 個のタガログ語-日本語ペアを発声し、参加者はタガログ語に対応する日本語訳を出来るだけ多く覚えるように指示された。1 ペアあたり 5 秒の間隔で 2 度ずつ提示され、10 ペアがランダムな順序で連続で提示された。Single 条件では、1 人の講師エージェントが 2 度反復して発声し、Double 条件では 2 人の講師エージェントが 1 度ずつ発声した（図 1）。HMD とヘッドホンを外し、椅子に座ると、インターバルタスクとして紙に印刷された簡単な計算課題（例： $4 + 2 \div 2$ ）に 30 秒間取り組んだ。この計算課題は、リハーサル（頭の中で単語を反復すること）を抑制するために設けられたタスクであった。続いて、モニタを見ながら想起テストを受けた。想起テストでは、スピーカ（Bose Soundlink II）からタガログ語が音声で提示され、参加者は対応する日本語訳を口頭で回答するように指示された。1 項目あたりの回答時間は 5 秒間で、10 項目がランダムな順序で連続でテストされた。テスト中、モニタには残りの回答時間が表示されていた。その後、参加者はバーチャルエージェントの社会的存在感や不気味さなどを測定する質問紙 [11] に回答した。

条件を変えて、覚える、インターバル、テスト、質問紙のサイクルを繰り返し実施した。最後に、参加者は SSQ と、VR 体験の主観的没入感を評価する Igroup Presence Questionnaire に回答した。その後、記憶課題についてのインタビュー（例：「既知の単語はありましたか？」）とバーチャルエージェントに関するインタビュー（例：「1 人の場合と 2 人の場合で、体験の違いはありましたか？」）に回答し、謝礼を受け取り解散した。実験は 1 時間程度であった。

3. 結果および考察

予備的調査であったことと紙面の都合から、以下では主要な結果として、記憶課題の正答率と不気味さの主観評価のみについて報告する。また、解析では、統計検定は実施せず、必要に応じて効果量を報告することとした。

3.1 正答率

1 名の参加者を除き、事前に知っている単語があったと回答した参加者はいなかった。1 名は、kahon（箱）という単語を知っていたと回答したため、先行研究 [6] を参考に、残りの 9 単語の回答のみから正答率を計算し、解析に含めた。

図 2 に正答率を示す。参加者の半数 (1, 4, 5, 6) は Double 条件優位で、半数 (2, 3, 7, 8) は Single 条件優位であった。参加者によって大きく傾向が異なったことから、独立変数であるエージェントの数以外に正答率に寄与した要因があった可能性が考えられた。以下では、順序効果の可能性に着目し、探索的な解析を実施した。

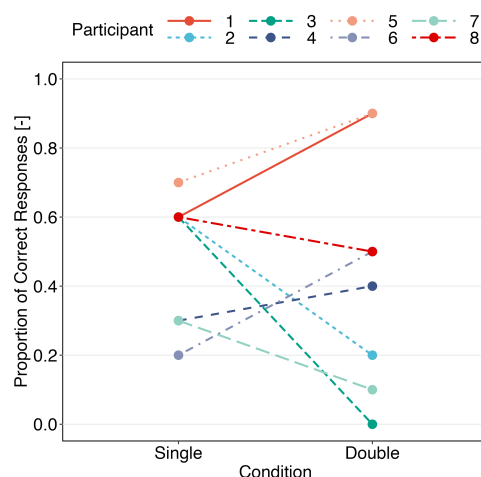


図 2: 正答率。参加者によって傾向が異なっていた。

図 3 に試行ごと（1 試行目、2 試行目）の正答率を示す。1 試行目では、Single 条件が Double 条件に比べて正答率が高かった (Cohen's $d = 1.20$)。他方、2 試行目では Double 条件が Single 条件に比べて正答率が高かった (Cohen's $d = 0.65$)。この結果は、タスクの認知負荷と Social Facilitation / Inhibition の関係 [12] に基づいて解釈できる可能性がある。タスクの認知負荷が低い場合、エージェントの存在はタスクパフォーマンスを向上する (Social Facilitation)。他方、認知負荷が高い場合、効果が逆転し、エージェントの存在はタスクパフォーマンスを阻害する (Social Inhibition)。同様に、本研究においては、記憶課題に慣れておらず認知負荷が高かった 1 試行目では 2 人のエージェントの存在が正答率にネガティブに作用したと考えられる。他方、記憶課題に慣れて認知負荷が低くなった 2 試行目では 2 人のエージェントの存在が正答率にポジティブに作用した可能性がある。この考察を踏まえ、今後の実験では、順序効果を排除するために、事前に記憶課題に慣れさせる試行を導入することが考えられる。または、記憶課題の難易度を 2 種類（簡単、難しい）を用いた検証によって、認知負荷と得られる効果の関係について洞察が得られるかもしれない。

3.2 不気味さ

図 4 に不気味さの主観評価を示す。Double 条件が Single 条件よりもスコアが高い傾向が見られた (Cohen's $r = 0.29$)。この不気味さは記憶成績に影響する可能性があるため、今後の実験でも測定し、その影響を注意深く観察する。また、不気味さを回避するならば、相異なる外見のエージェントや、ソーシャルロボットのように同一の外見をしていることに違和感のないエージェントを採用することも考えられる。

3.3 参加者からのフィードバック

講師エージェント 2 人の設計指針について、8 名中 7 名は異なる外見を採用する方法について述べた。具体的には、記憶課題の性質に合わせた外国人と日本人の組み合わせ、父母のアナログとしての男女の組み合わせ、服装の色違いなどが候補して挙げられた。特に、日本の英語教育においては

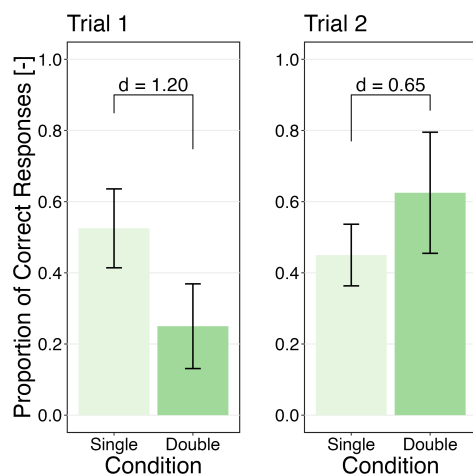


図 3: 試行ごとの正答率。1 試行目と 2 試行目では、異なる傾向が見られた。

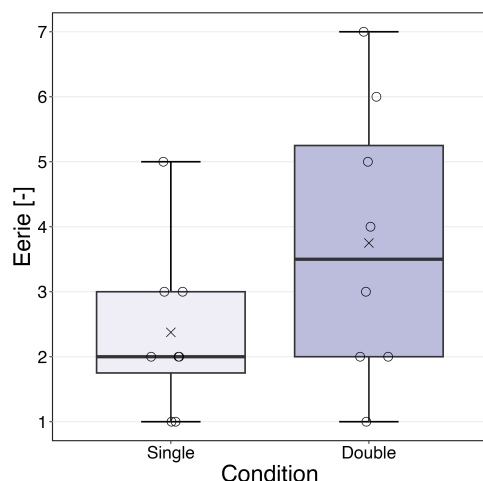


図 4: 不気味さ。Double 条件では、不気味さが高かった。

ネイティブスピーカーと日本人の 2 名の講師による授業形態が馴染み深いという意見は興味深かった。また、2 人が同じ声であったことについては、外見による違和感のほうが大きく、気にならなかったという参加者が多かった。ただし、異なる外見で同じ声だった場合には違和感があるだろうという指摘もあった。加えて、講師 2 名という状況が日常生活であまりないという意見や、講師 1 名と同級生 1 名のような使い方もアイデアとして挙げられた。また、2 人のエージェント同士のインタラクションがあると良いというコメントもあった。記憶成績に対して、これらの設計がどのように寄与するかは未解明であり、今後調査が期待される。

4. おわりに

本研究は、VR における講師エージェントを用いた学習支援システムの効果を向上するため、エージェントの数の効果に着目し、1 人の講師エージェントを用いる Single 条件と 2 人の講師エージェントを用いる Double 条件を参加者内比較した。8 名を対象とした予備的調査の結果、期待さ

れる効果はタスクの順序や認知負荷によって影響を受ける可能性があること示唆された。記憶課題の練習試行の導入や、記憶課題の難易度の操作、参加者内計画ではなく参加者間計画を採用するなど、引き続き調査を続ける。

謝辞 本研究の一部は、JST ムーンショット型研究開発事業 (JPMJMS2013) の支援を受けた。

参考文献

- [1] A. L. Baylor and Y. Kim: Pedagogical Agent Design: The Impact of Agent Realism, Gender, Ethnicity, and Instructional Role, *Intelligent Tutoring Systems*, 592–603, 2004.
- [2] T. Mizuho et al.: Investigating the Effects of Changing the Appearance of Screen-Based Avatars on Audience Memory, *ACM Symposium on Applied Perception*, 1–9, 2024.
- [3] M. Guimarães et al.: The Impact of Virtual Reality in the Social Presence of a Virtual Agent, *Proc. of the 20th ACM International Conference on Intelligent Virtual Agents*, 1–8, 2020.
- [4] G. B. Petersen et al.: Pedagogical Agents in Educational VR: An in the Wild Study, *Proc. of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–12, 2021.
- [5] G. Makransky et al.: A gender matching effect in learning with pedagogical agents in an immersive virtual reality science simulation, *Journal of Computer Assisted Learning*, 35(3), 349–358, 2019.
- [6] T. Mizuho et al.: Effects of Social Contextual Variation Using Partner Avatars on Memory Acquisition and Retention, *arXiv*, 2025.
- [7] M. Shiomi et al.: Two is better than one: verification of the effect of praise from two robots on pre-school children's learning time, *Advanced Robotics*, 35(19), 1132–1144, 2021.
- [8] M. Shiomi et al.: Two is better than one: Social rewards from two agents enhance offline improvements in motor skills more than single agent, *PLOS ONE*, 15(11), e0240622, 2020.
- [9] F. Yonemitsu et al.: The clone devaluation effect: A new uncanny phenomenon concerning facial identity, *PLOS ONE*, 16(7), e0254396, 2021.
- [10] M. Gonzalez-Franco et al.: The Rocketbox Library and the Utility of Freely Available Rigged Avatars, *Frontiers in Virtual Reality*, 1:561558, 2020.
- [11] K. Zibrek et al.: Is Photorealism Important for Perception of Expressive Virtual Humans in Virtual Reality?, *ACM Transactions on Applied Perception*, 16(3), 14, 2019.
- [12] M. R. Miller et al.: Social interaction in augmented reality, *PLOS ONE*, 14(5), e0216290, 2019.