

商品紹介動画におけるエージェント表現と ユーザ視線動作に関する研究

近藤 千紗^{1,†1} 佐久間 洋司^{2,3} 土方 嘉徳¹

概要：SNS やソーシャルメディアで活動するインフルエンサーは、自らのコンテンツで商品やサービスの利用体験を共有し、フォロワーの消費行動に対して影響を与える。インフルエンサーは動画ベースのメディアで商品紹介を行うことが増え、企業から商品紹介の依頼を受ける人もいる。また、インフルエンサーの見た目は多様化しており、エージェント表現の違いによって、ユーザが受ける情報量には差があると考えられる。そこで、本研究では商品紹介動画のエージェント表現の違いによって、広告効果（動画視聴後の購買意欲、商品に対する理解、ブランドへの態度）と視聴者の視線動作に与える影響が異なるのかを調べる。被験者は人間の出演する動画、アバターの出演する動画、デフォルトアイコンの画像が表示された動画うちのいずれか1本の動画を視聴し、視線計測装置によって動画視聴時の視線が取得された。その後、被験者はアンケートに回答した。実験の結果、3種類のエージェント表現で広告効果に差はなかったが、エージェントの身体的魅力と好感度と信頼性に差があることが分かった。また、人間とアバターの動画では、動画視聴中の視線の対象が異なることが分かった。

キーワード：YouTuber, VTuber, アバター, アイコン, 商品紹介動画, 視線計測, 被験者実験

1. はじめに

インフルエンサーはSNS やソーシャルメディアを通じて、他人の行動に対して影響を与える。彼らは、スポンサーの有無にかかわらず、自らのコンテンツで商品やサービスの利用体験を共有する。フォロワーはインフルエンサーと自身との間に類似性を見出し、インフルエンサーを信頼すると、従来の有名人に紹介されるより、商品を購入する意欲が高まる可能性があることが分かっている [1]。そのため、現在の企業の広報活動は、テレビや新聞といったマスメディアにおける広告だけではなく、インフルエンサーへの商品紹介の委託を行うことも増えている。

これまで、インフルエンサーによる商品紹介はテキストや画像を中心としたメディアで行われていたが、商品の利用方法をよりリアルに伝えることができるため、YouTube

や TikTok 等の動画ベースのメディアを用いる人が増加している。動画ベースのメディアは、視覚的な情報と音声による情報で構成されている。一方で、Clubhouse や Discord 等の音声ベースのメディアでは視覚的な情報はアイコンで表現されるもののみであり、メインとなる情報は音声で表現されている。動画ベースのメディアの中でも、YouTube におけるインフルエンサーは YouTuber と呼ばれている。YouTuber とユーザアイコンのみでは、ユーザの受け取る視覚的な情報量に大きな差がある。

また、YouTuber より視覚的な情報量が少ないインフルエンサーとして、バーチャル YouTuber (以降、VTuber と呼称) が考えられる。VTuber とは、3D モデル、2D モデル、または 2D のイラストを使って動画作成・投稿を行う YouTuber のことである。VTuber は見た目や声のカスタマイズ性が高く、自らの実際の見え方や声を気にせず動画に出演することができるため、多くの視聴者から人気を集めている。

YouTuber と同様、VTuber も動画内で商品紹介を行うことがあるが、エージェント表現の違いによる視覚的な情報量の差により、商品紹介動画の広告効果に差が生まれるかもしれないと考えた。最も情報量の多いエージェント表現として YouTuber を、次点で VTuber を、最後にユーザアイコンを挙げる。これら3つのエージェント表現の違い

¹ 関西学院大学商学部
School of Business Administration, Kwansai Gakuin University
² 大阪大学グローバルイニシアティブ機構
Center for Global Initiatives, Osaka University
³ 東京大学大学院総合文化研究科
Graduate School of Arts and Sciences, The University of Tokyo
^{†1} 現在、奈良先端科学技術大学院大学先端科学技術研究科
Presently with Graduate School of Advanced Science and Technology, Nara Institute of Science and Technology

によって、広告効果に差が出るのかを明らかにする。

また、インフルエンサーの社会的魅力や身体的魅力、態度の同質性は、消費者のブランドへの認識や商品の購買意欲にプラスの影響を及ぼすことが分かっている [2]。エージェントへの印象に差がある場合、広告効果に影響を及ぼすかもしれないと考える。そこで、3つのエージェント表現の違いによって、エージェントへの印象に差が出るのかを明らかにする。

YouTuber の動画と VTuber の動画では、見た目の違いにより、被験者の視線の対象が異なる可能性があると考えられる。YouTuber は実際の人間が出演するため、表情だけではなく、顔の表情や髪質感、服装の様々なデザインなど多くの情報が含まれる。一方で VTuber はアバターやイラストを用いるため、セルルックな見た目をしており、顔の表情にも制限がある。そこで、視線計測装置を使用して被験者の視線を取得し、人間条件とアバター条件のグループの被験者の視線の違いがあるか、被験者の視線の対象と広告効果には関連があるのかを明らかにする。

これらの目的より、次の4つの RQ (Research Question) を設定した。本研究での広告効果とは、動画視聴後の購買意欲、商品に対する理解、ブランドへの態度のことを指す。

- RQ1: 3種類のエージェント表現によって、広告効果に差が生じるか?
- RQ2: 3種類のエージェント表現によって、エージェントへの印象は異なるか?
- RQ3: 2種類のエージェント表現（人間とアバター）によって、動画視聴中の視線の対象が異なるか?
- RQ4: 動画視聴中に視線の対象と広告効果に関連はあるか?

2. 理論的背景

コンピュータで生成されたユーザの視覚的表現であるアバターは、仮想環境、SNS、ビデオゲーム、広告、映画といった幅広い分野で利用されている。e コマースサイトや仮想店舗環境でのアバターの存在が、購買意欲と商品に対する態度を向上させることが分かっている [3][4]。また、紹介する製品に対して被験者の製品カテゴリの知識が少ない場合、アバターが出演する広告より人間が出演する広告の方が広告への態度と購買意欲が高く、製品カテゴリの知識が多い場合は、人間とアバターの広告で有意差は見られなかった [5]。

アバターの擬人化はアバターに対する認識に影響し、より擬人化されたアバターはより魅力的で信頼性が高いと認識され、人々はそのようなアバターを選ぶ傾向があった [6]。エージェントの表現形式（非人間、人間、漫画的な人間）は、提示されたエージェントの画像の肉体的魅力に

関する不確実性に影響を与えることが分かっている。非人間の不確実性が最も高く、次点で漫画的な人間、最後に人間の写真の順であった [7]。これらの関連研究より、次の仮説 1 と仮説 2 を立てた。

- 仮説 1: 視覚的な情報量の多い人間による表現は、視覚的な情報量が少ないアバターによる表現やアイコン表現に比べて、広告効果が高くなる
- 仮説 2: 視覚的な情報量の多い人間による表現は、視覚的な情報量が少ないアバターによる表現やアイコン表現に比べて、良い印象になる

アバターモデリングソフトウェアで人間の顔の画像とアバターの顔の写真のモーフを作成し、目、鼻、口への視覚的注意を各段階で総停留時間と総停留回数を比較したところ、人間に近づくと目への注意が高まり、鼻への注意が低下した。また、全体的を通して目、鼻、口の順で視覚的注意が高いことが分かった [8]。

広告に対する想起（理解）、広告に対する態度、ブランドに対する態度、購買意欲の4つの広告効果と要素（製品、ブランド、商品の推奨者）ごとの眼球運動の動作との間には関連がある可能性があることが分かっている [9]。これらの関連研究より、次の仮説 3 と仮説 4 を立てた。

- 仮説 3: 人間とアバターの動画では、動画視聴中の視聴者の視線の対象が異なる
- 仮説 4: 動画視聴中の視線の対象と広告効果には関連がある

3. 実験方法

3.1 実験概要

被験者に商品紹介動画を視聴し、アンケートに回答してもらった実験を行った。条件は、人間の動画を視聴するグループ、アバターの動画を視聴するグループ、デフォルトアイコンの画像が表示された動画を視聴するグループの3つである。人間の動画は、著者らの1人が出演し、カメラを用いて撮影を行った。人間の動画を撮影する際に同時に、事前にかから作成した前記出演者の見た目に近いデザインの3Dモデルを3D空間上で出演させ、3D空間上で撮影を行うことによってアバターの動画を撮影した。出演者は撮影前に慣性式のトラッキングスーツを着用し、全身の動きをトラッキングすることによって、人間とアバターの動きが全く同じになるようにした。また、デフォルトアイコンの画像を用いて、アイコンの動画を作成した。被験者は各条件で25人ずつ、3つのグループに分けられた。動画を視聴している間、視線計測装置によって被験者の視線を取得した。その後、被験者はアンケートに回答した。

3.2 動画素材

実験素材として、人間とアバターが商品紹介を行う2本の動画を作成した（図1の左上、右上）。3Dモデルはト

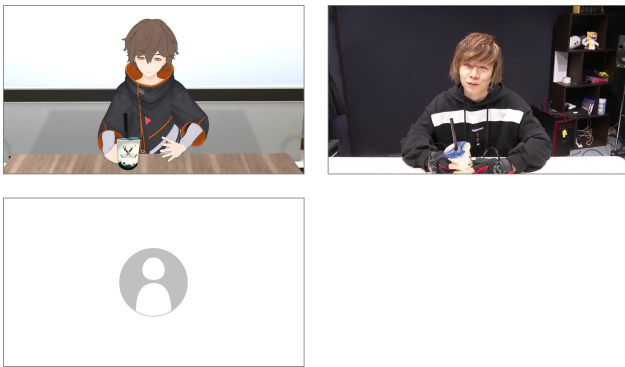


図 1 商品紹介動画の画面例 (左上: YouTuber, 右上: VTuber, 左下: アイコン)

ラッキングシステムのセンサーから取得した加速度に応じて動くため、人間とアバターの動画で出演者の動きや音声は全く同じである。

動画で紹介する商品ドメインには、タピオカドリンクを選んだ。動画内容は、出演者がティースタンド The Alley の 5 つの商品を試飲し評価を行うというものである。VTuber は 3D または 2D 空間でアバターを操作するため、現実世界の物体をアバターのいる空間にそのままの表現形式で持ち込むことができない。特に、食品や飲料物は実際に VTuber が消費することはできないため、表現に不自然が生じる。そこで、VTuber にとって不自然な状況でも購買意欲が高まるのか調べたいと考え、動画撮影時に日本で流行していたタピオカドリンクを紹介する商品に選択した。

動画は大まかに 4 つのシーンに分けられる。1 つ目は、出演者が企画内容について説明しているシーン (合計 70 秒)。2 つ目は、出演者が実際の商品の拡大した静止画を画面の左側で提示して商品の基本情報について説明しているシーン (合計 33 秒)。3 つ目は、出演者がタピオカドリンクを試飲し評価をしているシーン (合計 409 秒)。4 つ目は、場面切り替えのシーンである (合計 18 秒)。本研究では、被験者の動画視聴中に視線計測を行うため、動画にテキストがあると、被験者の注意がテキストを読むことに集中してしまうため、動画にテロップは付けなかった。

アイコンの動画は、動画と音源を記録・変換・再生することができる FFmpeg^{*1} を用いて、撮影した動画の音声のみを抽出し、デフォルトアイコンの画像 1 枚と抽出した音声データを合成して動画を作成した (図 1 の左下)。出演者のアイコンではなく、デフォルトアイコンにしたのは身体的な情報を減らし、意味のある情報がほとんどない条件を用意するためである。

3.3 アンケート基本設計

実験では主に、広告効果、エージェントへの印象、デモグラフィック指標、出演者や商品に対するユーザの知識と

経験についてアンケートで尋ねた。広告効果については、動画視聴後の購買意欲、商品に対する理解、ブランドへの態度を尋ねた。エージェントへの印象については、対人関係における社会的魅力と身体的魅力、対人コミュニケーションにおける知覚された態度の同質性、エージェントに対する好感度と信頼性を尋ねた。アンケートは、商品紹介動画を視聴した直後に行った。

3.4 アンケートの詳細

3.4.1 動画視聴後の広告効果

購買意欲は、実際に被験者にインフルエンサーの投稿を見せて購買意欲の程度を回答させた。この質問には、Stubb と Colliander が作成した 4 つの質問を日本語に訳して使用した [10]。質問項目は 7 点リッカート尺度で測定した。商品に対する理解は、独自に作成した 5 つの質問を使用し、7 点リッカート尺度で測定した。動画では出演者がタピオカドリンクを試飲し、見た目や味の特徴を説明していた。そこで、その話の内容をどの程度理解していたかについて尋ねた。ブランドへの態度は、MacKenzie と Lutz, Belch が作成した 3 つの質問を日本語に訳して使用し、7 点リッカート尺度で測定した [11]。動画視聴後の広告効果に関する各質問項目は表 1 に記載している。

3.4.2 エージェントへの印象

対人関係における社会的魅力は、McCroskey と McCain が作成した Social Attraction を測定する 6 つの質問を日本語に訳して使用し、5 点リッカート尺度で測定した [12]。対人関係における身体的魅力は、McCroskey と McCain が作成した Physical Attraction を測定する 6 つの質問を日本語に訳して使用し、5 点リッカート尺度で測定した [12]。対人コミュニケーションにおける知覚された態度の同質性は、McCroskey と Richmond, Daly が作成した Attitude Homophily を測定する 4 つの質問を日本語に訳して使用し、7 段階の SD 法で測定した [13]。エージェントに対する好感度は、独自に作成した 1 つの質問を使用し、7 点リッカート尺度で測定した。エージェントに対する信頼性は、Ohanian が作成した 4 つの質問を日本語に訳して使用し、7 段階の SD 法で測定した [14]。エージェントへの印象に関する各質問項目は表 1 に記載している。

3.4.3 デモグラフィック指標

デモグラフィック指標は、性別 (3 択から選択式)、年齢 (記述式)、居住地域 (3 択から選択式)、趣味 (複数選択式) の 4 つについて質問した。

3.4.4 出演者や商品に対するユーザの知識と経験

出演者を知っていた場合、紹介されていたブランドを知っていた場合、紹介されていたブランドのドリンクを飲

*1 <https://ffmpeg.org/>

表 1 アンケート項目

購買意欲
Q1 動画内で紹介されていた商品を買いたい。
Q2 他の情報源から動画内で紹介されていた商品についてさらに情報を集めると思う。
Q3 動画で紹介されていたブランドから商品を購入する可能性はどれくらいありますか？
Q4 もしあなたが将来対象ドメインの商品を購入する必要がある場合、動画で紹介されたブランドから購入する可能性はどれくらいありますか？
商品に対する理解
Q1 動画内で紹介されていたジアレいのタピオカドリンクを目の前で並べられたら、どれがどの商品かを当てる自信がある。
Q2 動画内で紹介されていたジアレいのタピオカドリンクを順に目隠して試飲したら、どれがどの商品かを当てる自信がある。
Q3 動画内で紹介されていたジアレいのタピオカドリンクと他社ブランドのタピオカドリンクを目の前で並べられたら、どちらがジアレいのタピオカドリンクかを当てる自信がある。
Q4 動画内で紹介されていたジアレいのタピオカドリンクが、どのような商品なのか友人に説明することができる。
Q5 動画内で紹介されていたタピオカドリンクのうちの任意の1つを、実際に店舗に行って、ネット等を検索しなくても注文できる。(店舗内にある情報は見てよい)
ブランドへの態度
Q1 私は動画内で紹介されていたブランドが好きです。
Q2 動画内で紹介されていたブランドへの私の印象は良いです。
Q3 動画内で紹介されていたブランドへの私の印象は肯定的です(ポジティブです)。
対人関係における社会的魅力
Q1 私は彼と友達になってもよいと思う。
Q2 彼と親しくおしゃべりしたい。
Q3 彼と会って話をするのは避けたいと思う。
Q4 彼が私の友達の輪の中に入ることはむずかしいだろう。
Q5 彼と私は決して個人的な友情を結ぶことはできないだろう。
Q6 彼と一緒にいると楽しいと思う。
対人関係における身体的魅力
Q1 彼はかなりハンサムだと思う。
Q2 彼はちょっと醜い。
Q3 彼はとてもセクシーな外見をしている。
Q4 彼の外見はとても魅力的だと思う。
Q5 私は彼の見た目が好きではない。
Q6 彼はあまり格好よくない。
対人コミュニケーションにおける知覚された態度の同質性
Q1 私に似ている - 私とは違う
Q2 私とはかけ離れていると思う - 私に近いと思う
Q3 私と同じように考えると思う - 私とは考え方が違う
Q4 私と同じようにふるまわない - 私と同じようにふるまう
エージェントに対する好感度
Q 動画の中で登場した人物(アバター)に対して、どの程度好感が持てましたか？
エージェントに対する信頼性
Q1 頼れない - 頼れる
Q2 誠実でない - 誠実な
Q3 正直でない - 正直な
Q4 信頼できない - 信頼できる

んでいた場合、紹介されていたブランドに行ったことがあった場合、該当する被験者とそうでない被験者では広告効果に差が出る可能性があると考えた。したがって、これらの回答を制御変数として得るために、以下の4つの質問を作成した。質問項目を以下にまとめる。

- Q1 この実験の前に、動画に登場した「佐久間洋司」という人物を知っていましたか？
- Q2 この実験の前に、動画で紹介されていた「ジアレイ」というブランドを知っていましたか？
- Q3 この実験の前に、ジアレいのドリンクを飲んだことがありますか？
- Q4 この実験の前に、ジアレいの店舗に行ったことがありますか？

3.5 オブジェクトフレームと視線計測データの分析方法

3.5.1 オブジェクトフレームとは

本研究では、被験者がどの対象物を注視しているのかを調べるため、オブジェクトフレームという矩形領域を動画のフレームごとに設定した。一般に、動画は静止画のコマ送り構成され、動画を構成する1枚1枚の静止画はフレームと呼ばれる。また、1枚の静止画は1フレーム(1コマ)と呼ばれる。1フレーム内において、オブジェクトが存在する場所(長方形の矩形範囲)のことを、オブジェクトフレームと本研究では呼ぶ。

実験で使用する動画には、実際に試飲するドリンク、右手、左手、顔、頭、身体、ドリンクの静止画像、背景、画面外の合計9個のオブジェクトが存在する。ただし、全てのオブジェクトが動画内で存在する時間を同じにするため、右手と静止画像のオブジェクトは分析では使用しない。視線の座標がどのオブジェクトフレームのX座標の範囲とY座標の範囲に入っているかを調べることで、被験者がどこを見ているのかを判断した。

3.5.2 オブジェクトフレームの設定方法

オブジェクトフレームは3つのステップにて設定した。1段階目に、FFmpegによって、8分50秒の人間とアバターの動画を1秒に1枚ずつフレームを切り出して530枚の画像を作成した(一定時間ごとにフレームを抽出した)。2段階目に、各8つのオブジェクトフレームのX座標とY座標の範囲を530フレーム分取得するため、学部の学生7人にアルバイトを依頼した。アルバイトの学生には多機能ペイントアプリPictBear^{*2}またはGIMP^{*3}を自身のPCにダウンロードし、自分の目で1枚ずつ担当のオブジェクトフレームのX座標とY座標の数値をスプレッドシートに書き込んでもらった。3段階目に、スプレッドシートに記入された数値で正確に動画内オブジェクトを囲っているか確

*2 <https://www.fenrir-inc.com/jp/pictbear/>

*3 <https://www.gimp.org/>

かめるため、Python の画像処理ライブラリである Pillow を使用して 530 枚の画像にオブジェクトフレームを書き込み、530 枚全てを目視でチェックした。一部の数値を修正し、530 フレーム分のオブジェクトフレームの X 座標と Y 座標の範囲を確定した。

3.5.3 視線計測データの分析方法

全シーンの構図を同じにするために、動画内で分析対象のオブジェクトが全て表示されているシーンのみを使用した。具体的には 3.2 で説明したシーン 3 の出演者がタピオカドリンクを試飲し評価をしているシーンのみを用いた (409 秒)。

実験では、60Hz のサンプリングレートで視線を検出する Tobii Pro Nano という視線計測装置を使用した。使用したモニターは 23 インチであった。この視線計測装置はコンピュータのモニター下部に取り付けられた。視線計測データの取得には、生体計測解析ソフトウェアである Tobii Pro Labo を使用した。視線計測データは Excel 形式で取得した。

今回の実験では視線の生データではなく、停留時の視線のデータを使用した。停留とは、視覚システムが見ているものに関する詳細な情報を取り込むことができるように、中心窩の視覚を保持したまま、目が比較的静止している期間のことを指す。停留の判定は、Tobii Pro Labo の Tobii I-VT Filter により自動的に判定された [15][16]。このフィルターによって、1 秒あたりの視覚度 ($^{\circ}/s$) が特定の閾値を超える場合にはサッカードとして分類され、閾値を下回ると停留の一部と分類される。サッカードとは、中心窩をある点から別の点へ急速に移動させるために用いられる眼球運動の一種である。

RQ3 と 4 ではどのオブジェクトフレームを注視しているかを調べる必要があるため、オブジェクトフレームごとに総停留時間を求める。使用するデータは以下の 3 つである。1 つ目に、Tobii I-VT Filter でフレームごとに出力された視線の状態 (停留かサッカード、未分類、瞳孔の未検出) の判定のデータ。2 つ目に、1 回の停留ごとに、停留期間中の予測された中心の注視座標のデータ。3 つ目に、1 フレームごとの時間のデータである。これらを用いて、以下の通りに各オブジェクトフレームの総停留時間を求めた。まず、事前にフレームごとに、オブジェクトの矩形領域であるオブジェクトフレームの左上の座標と右下の座標を設定した。次に、各フレームで視線の停留の座標 (X 座標と Y 座標) が得られているので、その座標がそのフレームのオブジェクトフレームの座標範囲に入っているかどうかを見て、視線がどのオブジェクトフレームに停留しているかを取得した。また、オブジェクトフレーム番号を用いて、各オブジェクトフレームの総停留時間と総停留回数を算出した。総停留時間は、オブジェクトフレームフレーム

番号ごとに 1 フレームごとの時間を合算して算出した。総停留回数は各オブジェクトフレーム番号の合計の数から算出した。

3.6 実験の実施

本研究は、関西学院大学の人を対象とする行動学系研究倫理委員会によって承認された。被験者の募集は、LINE と Instagram, メールを介して行われ、近畿圏に住む 19 歳から 61 歳までの 75 名のボランティア (男性 23 人, 女性 52 人, 平均年齢 27.48 歳) が実験に参加した。被験者は、男女比がほぼ均一になるように、3 つのグループ (人間, アバター, アイコン) に分けられた。具体的には、人間とアバター条件の参加者はそれぞれ男性 8 人, 女性 17 人, アイコン条件の参加者は、男性 7 人, 女性 18 人であった。実験に参加し、動画の視聴を行い、アンケートに回答した被験者には、1000 円の謝礼を支払った。動画視聴前に、視線計測装置のキャリブレーションを行うため、被験者は画面上の動く点を目で追った。キャリブレーション終了後、動画が再生された。動画の視聴が終わった後、被験者は本アンケートに回答した。全てのアンケートはオンラインでのアンケート作成ツールである Survey Monkey で作成された。

本研究では被験者からアンケートデータと視線計測データを取得した。アンケートデータの分析には 75 人分全てを使用する。一方で、視線計測データは 50 人分のデータのうち、総停留時間のヒストグラムを作成したところ、総停留時間が 100 秒以上 250 秒未満の範囲に一部外れ値が存在することが分かった。そこで、サンプリングレートが 70 % 未満または、総停留時間が 270 秒未満であった 5 人分のデータはサンプリングが不十分だとして分析に使用しないことにした。したがって、5 人のデータを削除し、45 人のデータのみを使用する。

4. 実験結果

4.1 RQ1: 3 種のエージェントによる広告効果の違い

3 種類のエージェント表現によって、動画視聴後の広告効果に差が生じるかを明らかにする。まず、動画視聴後の購買意欲、商品に対する理解、ブランドへの態度の 3 つの指標に対する回答の一貫性を確認するため、クロンバックの α 係数を確認した (表 2)。この結果より、いずれの条件においても概ね 0.7 以上の値となっており、回答の一貫性が確認された。

次に、3 つの指標を 3 種類のエージェント表現ごとに平均と標準偏差を算出し、また、データの分布の正規性を確かめるためにシャピロ・ウィルク検定を行い、検定結果の p 値を算出した (表 3)。シャピロ・ウィルク検定の結果、購買意欲では、人間の p 値は 0.927, アバターは 0.138, アイコンは 0.005 であった。アイコンのデータで正規性が確認されなかったため、3 群以上において平均値の差を検定

表 2 広告効果評価指標のクロンバックの α 係数

	購買意欲	商品理解	ブランドへの態度
Human	0.71	0.62	0.63
Avatar	0.77	0.68	0.83
Icon	0.73	0.72	0.85

表 3 表現形式と広告効果との関係に関する検定結果

	購買意欲		商品理解		ブランドへの態度	
	p=0.179		p=0.133		p=0.150	
	Mean	SD	Mean	SD	Mean	SD
Human	18.12	3.82	23.04	4.93	15.60	2.33
Avatar	17.36	4.55	20.28	5.33	14.04	3.75
Icon	19.24	4.42	22.48	4.95	15.92	3.07

表 4 エージェントの印象評価指標のクロンバックの α 係数

	社会的魅力	身体的魅力	同質性	信頼性
Human	0.81	0.76	0.78	0.74
Avatar	0.90	0.81	0.70	0.79
Icon	0.80	-	0.86	0.79

するノンパラメトリックな方法であるクラスカル・ウォリス検定を行った。その結果、 p 値は 0.179 であった。

商品に対する理解では、シャピロ・ウィルク検定の結果、人間の p 値は 0.412、アバターは 0.637、アイコンは 0.174 であった。正規性が確認されたため、パラメトリックな方法である一元配置分散分析を行った。その結果、 p 値は 0.133 であった。

ブランドへの態度では、シャピロ・ウィルク検定の結果、人間の p 値は 0.139、アバターは 0.419、アイコンは 0.020 であった。アイコンのデータで正規性が確認されなかったため、クラスカル・ウォリス検定を行った。その結果、 p 値は 0.150 であった。

以上の結果より、3種類のエージェントによって広告効果に差はなかった(表3参照)。したがって、「仮説1: 視覚的な情報量の多い人間による表現は、視覚的な情報量が少ないアバターによる表現やアイコン表現に比べて、広告効果が高くなる」は支持されなかった。

しかし、詳細な違いを見てみると、広告効果の平均値は、3つの指標全てでアバターが最も低く、購買意欲とブランドへの態度はアイコンが最も高く、商品に対する理解は人間が最も高かった。アイコンの購買意欲とブランドへの態度が高かったのは、視覚的な情報量が少なく処理が容易でメッセージが伝わりやすかったことが考えられる。一方、人間の商品に対する理解が高かったのは、アバター条件やアイコン条件とは異なり、実際に商品の見た目を動画で確認できることに加え、試飲時の顔の表情を見ながら説明を聞くことができるため、より話の内容が理解しやすかったからだと考える。

4.2 RQ2: 3種のエージェントによる印象の違い

3種類のエージェント表現によって、エージェントへの印象は異なるかを明らかにする。

まず、1項目からなる好感度を除いた、エージェントの社会的魅力、身体的魅力、態度の同質性、信頼性の4つの指標に対する回答の一貫性を確認するため、クロンバックの α 係数を確認した(表4)。この結果より、いずれの条件においても 0.7 以上の値となっており、回答の一貫性が確認された。

次に、5つの指標を3種類のエージェント表現ごとに平均と標準偏差、検定結果の p 値を算出した(表5)。エージェントの社会的魅力では、シャピロ・ウィルク検定の結果、人間の p 値は 0.876、アバターは 0.036、アイコンは 0.698 であった。アバターのデータで正規性が確認されなかったため、クラスカル・ウォリス検定を行った。その結果、 p 値は 0.109 であった。

エージェントへの態度の同質性では、シャピロ・ウィルク検定の結果、人間の p 値は 0.005、アバターは 0.064、アイコンは 0.187 であった。人間のデータで正規性が確認されなかったため、クラスカル・ウォリス検定を行った。その結果、 p 値は 0.330 であった。

エージェントの好感度では、シャピロ・ウィルク検定の結果、人間の p 値は 0.012、アバターは 0.010、アイコンは 0.029 であった。3つ全てのデータで正規性が確認されなかったため、クラスカル・ウォリス検定を行った。その結果、 p 値は 0.007 であった。

エージェントの信頼性では、シャピロ・ウィルク検定の結果、人間の p 値は 0.096、アバターは 0.919、アイコンは 0.008 であった。アイコンのデータで正規性が確認されなかったため、クラスカル・ウォリス検定を行った。その結果、 p 値は 0.009 であった。

エージェントの身体的魅力に関しては、人間とアバターのデータのみを比較した。なぜなら、アイコン条件では、動画に表示されるのはユーザアイコンのみで、エージェントの身体的魅力の質問に被験者が回答することが困難であると予測されたためである。エージェントの身体的魅力では、シャピロ・ウィルク検定の結果、人間の p 値は 0.624、アバターは 0.109 であった。正規性が確認されたため、2つの群の平均値の差を検定するパラメトリックな方法である t 検定を行った。 t 検定の結果、人間よりアバターの方が平均値が高く、 p 値は 0.035 であった。5つの分析の結果、エージェント間で、エージェントの社会的魅力とエージェントへの態度の同質性は有意差がなく、一方で、エージェントの身体的魅力と好感度と信頼性は有意差があった。

さらに、どのエージェントで差が出たのかを調べるため、エージェントの好感度と信頼性のデータで多重比較を行った(表6)。どちらもデータは3群で非正規分布であったため、モンテカルロ法によるスティール・ドゥワス検定を

表 5 表現形式とエージェントの印象との関係に関する検定結果

	社会的魅力		身体的魅力		同質性		好感度		信頼性	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
	p=0.109		p=0.035		p=0.330		p=0.007		p=0.009	
Human	19.00	3.81	17.76	3.48	12.24	4.44	4.08	1.15	16.80	4.05
Avatar	17.80	4.88	20.12	4.20	12.72	3.86	4.52	1.26	17.88	4.49
Icon	20.72	3.16	-	-	14.64	4.23	5.08	1.15	19.52	4.00

表 6 多重比較の検定結果

	好感度	信頼性
Human & Avatar	0.275	0.370
Human & Icon	0.004	0.005
Avatar & Icon	0.269	0.184

表 7 各オブジェクトフレームの平均総停留時間（秒）と検定結果

	ドリンク	左手	顔	頭	身体	背景	モニター外	モニター内の合計時間
	p=0.480	p=0.478	p=0.086	p=0.991	p=0.004	p<.001	p=0.555	p=0.634
Human	78.11	3.18	195.61	5.14	31.57	11.50	1.85	326.56
Avatar	86.02	4.04	171.10	6.96	46.58	4.32	3.51	323.05

行った。検定の結果、エージェントの好感度は人間とアイコンの組み合わせで有意差があり、p 値は 0.004 であった。エージェントの信頼性も同様に、人間とアイコンの組み合わせで有意差があり、p 値は 0.005 であった。また、エージェントの好感度と信頼性の平均値を見ると、人間よりアイコンの方が平均値が高いことが分かる（表 5 参照）。加えて、エージェントの印象に関するほとんどの質問では、アイコン、アバター、人間の順で平均値が高かった。したがって、「仮説 2：視覚的な情報量の多い人間による表現は、視覚的な情報量が少ないアバターによる表現やアイコン表現に比べて、良い印象になる」は支持されず、仮説と逆の結果であった。全体的に、視覚的情報量が少ないエージェント表現の方がエージェントに対する印象が良かったのは、視覚的情報量が少なくなることで音声に集中し、被験者がエージェントへのイメージを自分の中で膨らませることができたからかもしれない。

4.3 RQ3: 人間とアバターでの視線の対象の違い

2 種類のエージェント表現（人間とアバター）によって、動画視聴中の視線の対象が異なるかを明らかにする。動画視聴時の被験者の視線を人とアバターで比較するため、被験者の各オブジェクトフレームの平均総停留時間を算出した（表 7）。また、オブジェクトフレームごとにシャピロ・ウィルク検定と F 検定を行い、結果に応じて検定を行った。ドリンクでは、人間もアバターもデータは正規性があり、分散が等質であったため、t 検定を行った。p 値は 0.480 であった。左手では、人間もアバターもデータは正規性がなかったため、ウィルコクソンの順位和検定を行った。p 値は 0.478 であった。顔では、人間もアバターもデータは正規性があり、分散が等質であったため、t 検定を行った。p 値は 0.086 であった。頭では、人間もアバターもデータは正規性がなかったため、ウィルコクソンの順位和検定を行った。p 値は 0.991 であった。身体では、人間もアバターもデータは正規性がなかったため、ウィルコクソンの順位和検定を行った。p 値は 0.004 であった。背景では、人間

もアバターもデータは正規性がなかったため、ウィルコクソンの順位和検定を行った。p 値<.001 であった。モニター外では、人間もアバターもデータは正規性がなかったため、ウィルコクソンの順位和検定を行った。p 値は 0.555 であった。モニター内の合計時間では、人間もアバターもデータは正規性があり、分散が等質であったため、t 検定を行った。p 値は 0.634 であった。検定の結果、身体と背景では人間とアバターで有意差があり、その他のオブジェクトフレームでは差がないことが分かった。したがって、「仮説 3：人間とアバターの動画では、動画視聴中の視聴者の視線の対象が異なる」は支持された。

実験後に行った簡単なインタビューでは、以下のような意見があった。1 つ目に、人間の顔を多く見ていたのは、ドリンクを飲んだ時の表情やリアクションが気になっていたからである。2 つ目に、アバターの身体を多く見ていたのは、アバターの不自然な動きが気になっていたからである。3 つ目に、人間の背景を多く見ていたのは、背景にある棚が気になったからである。これらの意見は、人間とアバターで各オブジェクトフレームの平均総停留時間に差があった原因の一部であると考えられる。

4.4 RQ4: 視線の対象と広告効果との関連

動画視聴中の視線の対象と広告効果に関連はあるかを明らかにする。まず、分布の正規性を確かめるため、被験者ごとの各オブジェクトフレームの総停留時間のデータに対してシャピロ・ウィルク検定を行った。その結果、ドリンクの p 値は 0.127、顔の p 値は 0.338 で正規性が確認され、その他のオブジェクトフレームでは正規性が確認されなかった。そこで、被験者ごとの各オブジェクトフレームの総停留時間と広告効果に関連するアンケート項目のデータで、ノンパラメトリックな手法である Spearman の順位相関係数を算出し（表 8）、Spearman の無相関検定を行った。各オブジェクトフレームの平均総停留時間と広告効果（購買意欲、商品に対する理解、ブランドへの態度）の相関係数はいずれも低く、p 値はすべて 0.05 を上回った。これ

表 8 各オブジェクトフレームの平均総停留時間と広告効果の相関係数

	ドリンク	左手	顔	頭	身体	背景	モニター外	モニター内の合計時間
購買意欲	0.12	0.24	0.03	0.01	-0.06	0.02	0.09	0.12
商品に対する理解	-0.10	0.07	0.19	0.24	-0.24	0.17	0.01	0.14
ブランドへの態度	0.16	0.22	0.08	-0.16	-0.24	0.11	0.12	0.20

ら3つの結果から、広告効果と各オブジェクトフレームの平均総停留時間に相関がないことが分かり、「仮説4：動画視聴中の視線の対象と広告効果には関連がある」は支持されなかった。

5. おわりに

本稿では、商品紹介動画のエージェント表現の違いによって、広告効果と視聴者の視線動作に与える影響が異なるのかについて調査を行った。実験の結果、4つのことが分かった。1つ目に、3種類のエージェント表現によって、広告効果に差はないことが分かった。2つ目に、3種類のエージェント表現によって、エージェントの肉体的魅力と好感度と信頼性に差があることが分かった。3つ目に、2種類のエージェント表現によって、動画視聴中に視線の対象は異なることが分かった。4つ目に、動画視聴中に視線の対象と広告効果に関連はないことが分かった。

本研究では、動画に出演する人物は1人で、アバターも1体のみを使い、出演者の個性やアバターのデザインは1パターンのみとなった。より今回の結果を一般化するため、出演者の性別を複数用意し、様々な個性やアバターのデザインを用意して実験を行うとさらなる発見がある可能性が考えられる。また、紹介する商品はタピオカドリンクのみであった。タピオカドリンクをはじめとした食品ドメインは消費者は低関与になりやすい[17]。そのため、商品ドメインを変更した場合、結果に違いが出る可能性がある。

また、本研究では学習効果を避けるため、動画の視聴は1回のみであった。動画を複数回視聴することにより、結果が異なる可能性は考えられる。

謝辞

本研究は、孫正義育英財団の支援を受け、実施されました。本研究の一部は、JST CREST JPMJCR20D4の支援を受けたものです。また、YouTuberとVTuberの動画作成に協力してくださったActiv8社に感謝いたします。

参考文献

- Schouten, A. P., Janssen, L. and Verspaget, M.: Celebrity vs. Influencer endorsements in advertising: the role of identification, credibility, and Product-Endorser fit, *International journal of advertising*, Vol. 39, No. 2, pp. 258–281 (2020).
- Lee, J. E. and Watkins, B.: YouTube vloggers' influence on consumer luxury brand perceptions and intentions, *Journal of Business Research*, Vol. 69, No. 12, pp. 5753–5760 (2016).
- Holzwarth, M., Janiszewski, C. and Neumann, M. M.: The influence of avatars on online consumer shopping behavior, *Journal of marketing*, Vol. 70, No. 4, pp. 19–36 (2006).
- Jin, S.-A. A. and Bolebruch, J.: Avatar-based advertising in Second Life: The role of presence and attractiveness of virtual spokespersons, *Journal of Interactive Advertising*, Vol. 10, No. 1, pp. 51–60 (2009).
- Gammoh, B. S., Jiménez, F. R. and Wergin, R.: Consumer attitudes toward human-like avatars in advertisements: The effect of category knowledge and imagery, *International Journal of Electronic Commerce*, Vol. 22, No. 3, pp. 325–348 (2018).
- Nowak, K. L. and Rauh, C.: The influence of the avatar on online perceptions of anthropomorphism, androgyny, credibility, homophily, and attraction, *Journal of Computer-Mediated Communication*, Vol. 11, No. 1, pp. 153–178 (2005).
- Westerman, D., Tamborini, R. and Bowman, N. D.: The effects of static avatars on impression formation across different contexts on social networking sites, *Computers in Human Behavior*, Vol. 53, pp. 111–117 (2015).
- Cheetham, M., Pavlovic, I., Jordan, N., Suter, P. and Jancke, L.: Category processing and the human likeness dimension of the uncanny valley hypothesis: eye-tracking data, *Frontiers in psychology*, Vol. 4, p. 108 (2013).
- Zhang, X. and Yuan, S.-M.: An eye tracking analysis for video advertising: Relationship between advertisement elements and effectiveness, *IEEE access*, Vol. 6, pp. 10699–10707 (2018).
- Stubb, C. and Colliander, J.: “This is not sponsored content” –The effects of impartiality disclosure and e-commerce landing pages on consumer responses to social media influencer posts, *Computers in Human Behavior*, Vol. 98, pp. 210–222 (2019).
- MacKenzie, S. B., Lutz, R. J. and Belch, G. E.: The role of attitude toward the ad as a mediator of advertising effectiveness: A test of competing explanations, *Journal of marketing research*, Vol. 23, No. 2, pp. 130–143 (1986).
- McCroskey, J. C. and McCain, T. A.: The measurement of interpersonal attraction (1974).
- McCroskey, J. C., Richmond, V. P. and Daly, J. A.: The development of a measure of perceived homophily in interpersonal communication, *Human Communication Research*, Vol. 1, No. 4, pp. 323–332 (1975).
- Ohanian, R.: Construction and validation of a scale to measure celebrity endorsers' perceived expertise, trustworthiness, and attractiveness, *Journal of advertising*, Vol. 19, No. 3, pp. 39–52 (1990).
- Olsen, A.: The Tobii I-VT fixation filter, *Tobii Technology*, Vol. 21, pp. 4–19 (2012).
- Tobii Technology, A.: Determining the Tobii I-VT fixation filter's default values: method description and results discussion (2012).
- Zaichkowsky, J. L.: Conceptualizing involvement, *Journal of advertising*, Vol. 15, No. 2, pp. 4–34 (1986).