

VR 会議における他者との対面量と言語行動の 対応関係の分析

渡会 隆哉¹ 鎌田 光太郎² 角 康之¹ 王子洋² 由井 蘭 隆也² 高島 健太郎²

概要: 本研究は、VR 会議において見る/見られるといった対面量と参加者の発話にはどのような対応関係があるのかを調査する。VR 環境のような実際に対面していない状況では、参加者は発話している参加者を見るのかは分からない。この調査を行うことで、VR 会議での言語行動や対面量が活発、非活発になる原因を知ることができる可能性がある。これが分かれば、言語行動や対面量を活発にするような環境を整えることができ、会議の質を高めることにも繋がると考える。分析を行うために2つの環境を用意して実験を行い、環境の違いによって言語行動と対面量の対応関係には違いがあるのかを調査した。また、対面量と発話時間の対応関係を可視化した。これにより、2つの環境の分布には大きな差が見られることが分かった。壁提示環境では、参加者ごとで見ていた時間、見られていた時間、発話時間に大きな差は表れず、まとまった分布となっていた。机上提示環境では、参加者ごとで見ていた時間、見られていた時間、発話時間に差が表れることが分かった。これらのことより、壁提示環境ではまとまった行動が多くなり、そのため対面量と発話が均等にある可能性が示唆された。机上提示環境では、進行を行う参加者が現れ、発話に偏りが生じ、立ち位置によって対面量にも偏りが生じることが示唆された。

1. はじめに

本研究は、VR 会議において見る/見られるといった対面量と参加者の発話などは、どういった対応関係があるのかを調査した。

実世界の会議では、他の参加者に視線をを向けるなどといった非言語情報は、会議を行う上で重要な要素である。発話をしている参加者は、他の参加者を見ることで話をどう展開するかを伺う。また、話者を見ている参加者は、相手に視線を向けることで話を聞いているという態度を示すことができる。このように、非言語情報はコミュニケーションにおいて重要な情報である [1]。

また、会議において、発話などの言語行動も重要な要素である。発話をする量によっては、その参加者の印象は変化することが考えられる。発話量が大きくなればそれだけ会議に真摯に望んでいたことが伺え、発話量が小さければ、会議に対して参加意欲がさらに、発話をしている際は他の参加者の様子を伺いながら発話することで相手に話を振るなどの行動も行いやすくなる。このように、非言語情報と言語情報は密接な関係であり、発話行動を取ることによって非言語行動も促される。発話を多く行っている参加者は多くの視線を集めることが考えられる。しかし、VR(Virtual

Reality) 会議のような実際に参加者同士で対面しない状況でも言語行動と対面量には関係があるのかは分からない。音声だけ聞こえていれば会議は行えるため、話している参加者の方を見る必要はなく、そのような状況にある参加者は実際どういった行動を取るのかは分からない。対面していない分、話を聞いていることや話を伝えようとアピールするために真剣に相手の方を見る可能性も考えられる。

以前行った研究 [2] では、発話行動と社会的な手抜き [3] の認知を調査し、VR 環境の環境設定によっては言語行動は、参加者を評価する指標となり得ることが示唆された。もう1つ以前行った研究 [4] では、見る/見られるといった対面量に着目し、社会的な手抜きの認知や会議の質に対面量はあるのかを調査した。結果として、VR 環境の環境設定では対面量は参加者を評価する指標となり、会議の質とも相関関係があることが示唆された。しかし、これらの研究では言語行動と対面量の別々に着目していたため、それぞれがどういった対応関係になるかがわかっていない。

そこで本研究では、VR 会議において対面量と言語行動にはどういった対応関係があるのかを調査する。この関係性を明らかにすることができれば、VR 会議において参加者がどういった行動を取っているのかを明らかにすることができる。参加者がどういった行動を取っているのかを明らかにすることができれば、参加者の行動に合わせた会議

¹ 公立はこだて未来大学 システム情報科学研究科

² 北陸先端科学技術大学院大学 先端科学技術研究科

環境を作ることができ、会議を行いやすくすることに繋がる可能性がある。また、VR 会議での言語行動や対面量が活発、非活発になる原因を知ることができる可能性がある。これが分かれば、言語行動や対面量を活発にするような環境を整えることができ、会議の質を高めることにも繋がる。

2. 関連研究

2.1 グループワークでの言語行動と非言語行動

榎本らの研究 [5] では、日本語多人数会話における視線行動を分析し、話者交代と視線にはどういった関係があるのかを調査した。その結果、話者が視線を向けている参加者が次の話者になることが多いということが分かった。

このように、話者の交代というコミュニケーションを円滑に進める上で重要な工程にも、視線が関係している。このことから、発話という言語行動と見る/見られるといった対面量には密接な関係があると考えられる。そのため、本研究では VR 会議でも言語行動と対面量には関係があるのか、関係がないのかを分析する。

2.2 VR でのグループワークに関する研究

Wallace と Maryott の研究 [6] では、仮想環境のグループワークの質とアバターの外見には関係があるのかを調査した。異なる民族のアバターを用意し、グループワークに参加する上でどのアバターを利用したいかを調査した。結果として、参加者は自分と似た外見のアバターを好むことが分かった。また、他の参加者のアバターの見た目は、グループワークに対する意欲にも影響があることが分かった。このように、アバターの見た目は会議に対する意欲に関係する可能性がある。

そこで本研究は参加者のアバターは全て同じシンプルな見た目にし、参加者に余計な影響を及ぼさないようにした。

Han らの研究 [7] では、VR 環境のデザインが参加者に影響を与えるのかを調査した。仮想環境をパノラマの屋外仮想環境、パノラマの屋内仮想環境、広さに制限がある屋外仮想環境、広さに制限がある屋内仮想環境で実験をした。その結果、パノラマ環境のような広々とした空間の方が参加者同士で同期的な動きがあり、広さに制限のある仮想環境よりも存在感や喜び、楽しさなどが大きく感じられるということが分かった。このように、環境の違いによって参加者の感覚に影響を受ける可能性がある。

そこで本研究は、2つの仮想環境を用いて、環境の違いによって対面量や言語行動に影響があるのか、また、対面量と言語行動の対応関係は2つの環境で異なるのかを調査した。

3. 実験

3.1 会議環境

VR 会議を行うための仮想環境には Mozilla Hubs*1を用いた。VR 環境には図 1、図 2 と 2つの VR 環境を用いた。2つの VR 環境には、砂漠生き残り課題を説明するホワイトボード、テーブル、図 3 のようなアイテムの 3D モデルとアイテムに関する専門家の意見が書かれたポスター、水槽と小説が書かれたホワイトボードが配置された。テーブルとアイテムの 3D モデルはセットであり、これらの配置を変えることで、2つの環境を設定した。以下が2つの環境の違いである。

- **壁提示環境** (図 1) : アイテムとアイテムに関する説明情報を記載したポスターを部屋を取り囲むように周囲の壁に配置した環境。情報収集のために参加者が散策することが期待される。
- **机上提示環境** (図 2) : アイテムとアイテムの説明情報用紙を部屋の中心のテーブル上に並べ、密集させた環境。参加者全員がテーブル周辺に集まって議論を行うため、参加者同士の距離は近く、移動やグループの分裂は抑制されると思われる。

2つの VR 環境の違いにより、参加者の行動は影響を受けることが予想される。以前我々は、この VR 環境の違いは発話権取得失敗にどのような影響を与えるのかを調査した [8]。本研究では、VR 環境の違いによって、言語行動と対面量の対応関係は異なるのかを調査する。

VR 会議では図 4 のような、参加者に与える影響を最小限にするために白色のシンプルなアバターを使用した。胸のアルファベットと頭の上に浮かんでいるボードからどの参加者が識別することができる。また、発話を行うとアバターの上部が点灯する仕様になっている。さらに、顔の向きが分かるようなデザインのため、誰がどこを見ているのか概ね分かるようになっている。



図 1 壁提示環境の様子

*1 <https://hubs.mozilla.com/>

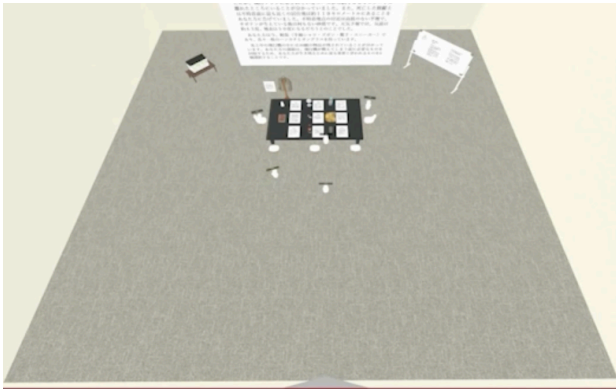


図 2 机上提示環境の様子

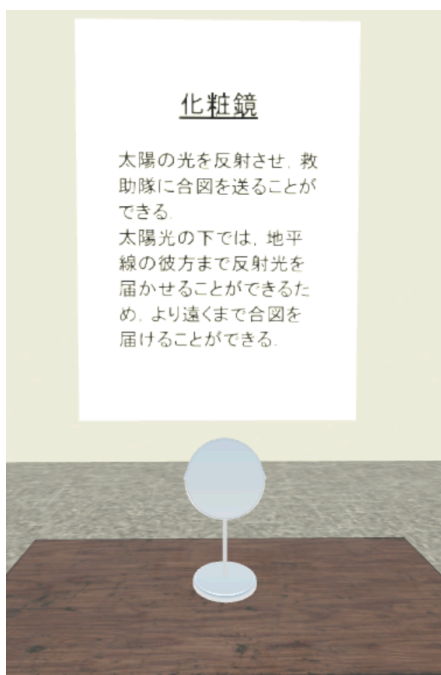


図 3 アイテムとポスターの例



図 4 VR 会議の様子（アバターは頭上のボードと胸の文字で識別可能）

3.2 実験に使用した課題

VR 会議には Lafferty ら [9] によって開発された砂漠生き残り課題を編集して使用した。この砂漠生き残り課題は、砂漠で遭難した際に、生存するために必要だと考えるアイテムの選択をする課題である。本研究では会議開始前と終了後、参加者にアンケートを実施して選んだアイテムを答えてもらった。本実験では、10 個のアイテムがあるうちの優先度が高いものを 5 個選択するという内容に変更した。

3.3 参加者

北陸先端科学技術大学院大学 (25 人) および、公立はこだて未来大学 (15 人) から計 40 人 (女性 8 人, 男性 32 人) を募集した。実験中に、1 人の辞退者が出たため、合計で 39 名となった。39 名の参加者が、8 グループ (4 名のグループが 1 つ, 5 名のグループが 7 つ) に分かれて VR 会議を実施した。

3.4 手続き

参加者には全員、Web カメラ付きのパーソナルコンピューター (PC) を用いて実験に参加するように求めた。また参加者には実験中、WebEx^{*2}を用いて画面共有をしてもらい、この様子を全て録画した。

1 回の VR 会議につき参加者は 5 人で行い、合計で 8 回行った。会議中には、基本的に自由に話してもらうが、個人が特定できる情報などは話さないように制限した。また、参加者には ID を割り振って、その ID を使用してもらった。

実験の始めに参加者には WebEx と Mozilla Hubs に接続してもらい、そこで課題の説明を行った。説明が終わった後、参加者には他の参加者とは相談をせず、15 分の時間を設けアンケートフォームで優先順位の高いアイテムを 5 個を選択してもらった。アンケートの回答後、40 分間の会議を開始した。会議終了後、15 分の時間を設け参加者にグループで決定した優先度の高いアイテム 5 個をアンケートフォームで回答してもらった。グループで決定したアイテムをアンケートフォームの回答後、15 分の時間を設け、自分は会議に貢献することができたかを評価する自己評価の項目が含まれたアンケートに回答するよう求めた。

4. 分析方法

4.1 対面量のデータ

実験で録画された、参加者の VR 会議での一人称映像からその参加者がどの参加者をどれだけの時間見ていたのかをラベリングした。ラベリングには ELAN^{*3}を用いた。参加者の 1 人称視点映像に、他の参加者が映り込んだ時点をラベルの始点として、映り込んだ参加者が視界から外れた

*2 <https://www.webex.com/ja/index.html>

*3 <https://archive.mpi.nl/tla/elan>

時点をラベルの終点とし、他の参加者を「見ていた」状態をラベリングした。複数の参加者が同じ時間に映り込んでいた場合は、複数人を見ていた状態としてラベリングを行った。なお、会議は各5人（1グループのみ4人）で行われていたため、最小で0人（誰も見ていなかった状態）から最大で4人（4人のグループのみ3人）を見ていた状態がある。他の参加者が視界に映り込んでいたかどうかを判別しラベリングを行ったため、参加者が実際に他の参加者を知覚し見ていたかどうかの判別はしていない、よって正確には、このラベルは参加者が他の参加者を見ていたという状態の近似である。

このラベリングデータからは以下のようなデータが得られる。

- 「見ていた/見られていた」時間（秒・ミリ秒）

ある参加者が他の参加者を「見ていた」時間は、ラベリングデータをそのまま参照することで得ることができる。また、「見ていた」時間は見られていた参加者側からすると「見られていた」時間であるため、「見られていた」時間もこのラベリングデータから得ることができる。

このデータから、参加者はどれくらい他の参加者を見ていたか、逆にどれくらい見られていたかを知ることができる。

4.2 言語行動のデータ

この実験では、参加者の発話も録音していた。そのため、記録データを用いて Whisper^{*4}や Adobe Premiere Pro^{*5}で書き起こしを行った。書き起こしたデータから以下のデータを得ることができた。

- それぞれの参加者の発話時間（秒・ミリ秒）

このデータから、参加者が会議中にどれくらいの発話を行っていたのかを明らかにすることができる。

4.3 得られたデータを用いて行った分析

言語行動と対面量の対応関係を理解するために以下の分析を行った。

見る/見られるといった対面量を x 軸と y 軸にとり、壁提示環境と机上提示環境2つの散布図を作成した。この散布図を見ることで、見る/見られるという対面量には環境によってどのような違いがあるのかを比較することができる。また、マーカーの大きさを発話時間の大きさと対応させる。そうすることで、対面量の分布と言語行動の分布を同時に理解することができる。また、発話時間が長い参加者は、見られている時間はどうなっているのかを確認することができる。この分布を見ることでそれぞれの環境での傾向やグループごとの傾向、特徴的な参加者を知ることができ、VR 会議での言語行動と対面量の対応関係を理解

し、こういったケースがあるのかを知る手がかりになると考える。

5. 結果と考察

5.1 対面量の分布と言語行動の対応の可視化

対面量の分布と言語行動の対応を図5に示す。図5は、x軸が見ていた時間（秒）、y軸が見られていた時間（秒）で構成され、それぞれの参加者がどういった非言語行動を行っていたかを視覚化した散布図である。x軸の中央線は見ていた時間の平均値、y軸の中央線は見られていた時間の平均値である。マーカーの大きさはそれぞれの参加者の発話時間と対応している。それぞれの発話時間には正規化を行っており、最小値0、最大値1となるようにスケールしている。左の散布図が壁提示環境、右の散布図が机上提示環境であり、それぞれのグループごとに色分けを行っている。

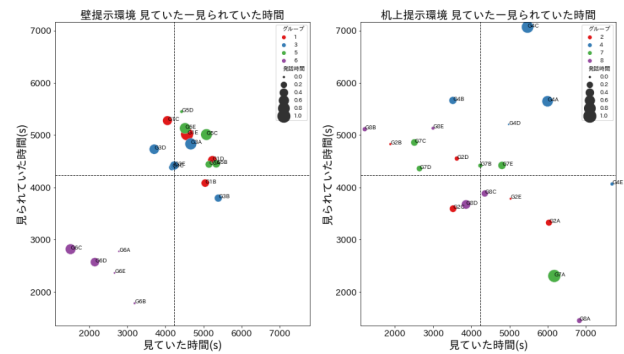


図5 対面量の分布と対応する言語行動の可視化

2つの散布図を比較すると、分布が大きく異なっていることが分かる。左の壁提示環境は一部を除き参加者同士の分布は固まっている。一方、右の机上提示環境は参加者同士で比較的まばらな分布となっている。

左の壁提示環境では、左下の部分に分布しているグループがある。このグループは会議が早く終わってしまったグループであり、見ていた時間も見られていた時間も短くなっている。その他の参加者は見ていた時間にも、見られていた時間にも大きな差はなく同程度の位置に集中している。また、グループごとでも、見ていた時間と見られていた時間に大きな差はなくまとまった分布となっている。見ていた時間も見られていた時間も長い参加者は、発話時間も長くなっている。さらに、参加者ごとで発話時間には大きな差はなく、参加者たちは同程度の発話時間で会議が行われていたことがわかる。

一方、右の机上提示環境は、まばらな配置となっており、見られていた時間の平均値あたりに分布している参加者は発話時間が短くなっている。発話時間が高くなっている参加者は、見られていた時間が短い場合と見られていた時間が長い場合の両極端の位置に見られた。また、グループの

*4 <https://openai.com/research/whisper>

*5 <https://www.adobe.com/jp/products/premiere.html>

中で参加者ごとに発話時間には大きな差があり、発話時間が長い参加者と発話時間が短い参加者が極端に分かれているグループが見られた。

5.2 壁提示環境での対面量の分布と言語行動の対応の考察

壁提示環境では、参加者ごとで見ていた時間と見られていた時間に大きな差はなく、まとまった分布となっていた。また、グループごとでもまとまった分布となっている。これは、グループごとでまとまって会議を行っていたことが考えられる。会議中に違う場所に移動する時も、まとまって動いていたため、見られていた時間にも見られていた時間も大きな差はなかったのではないかと考えられる。さらに、グループ全体でまとまって行動していたため、発話する機会も均等にあったことが考えられる。

また、見ていた時間と見られている時間は長い参加者は、発話時間も長いことが分かった。この環境では移動を行う機会が多いため、他の参加者の様子を伺うことも多くなる。そうした中で、発話をしている参加者を見る機会が増えるのは自然であり、発話をしている参加者は他の参加者の様子を伺うために他の参加者を見る機会が増えるというのも自然であるため、このような結果になるのは妥当であると考えられる。

このように、動きが促されるような環境では参加者は他の参加者の様子を伺いながら会議を進めるため、まとまった行動が多くなり、結果的に見ていた時間と見られていた時間、発話時間が同程度になることが示唆された。また、見る/見られるといった行動を見ることで、発話時間も知ることができるということも示唆された。

5.3 机上提示環境での対面量の分布と言語行動の対応の考察

机上提示環境では、参加者ごとで見ていた時間にも見られていた時間にも大きく差が見られた。また、発話時間は長い参加者と短い参加者の差が大きく表れた。これは、机上提示環境では立ち位置が一旦固定すると大きく変化することはあまりなく、基本は机側を見て会議に臨んでいたため、立ち位置によっては全く見られない位置になる場合もあり、見る時間も長くなる参加者もいれば、視界にあまり参加者が入らない位置だと見る時間も短くなる。よって、見る時間と見られる時間の差はそれぞれで大きくなったことが考えられる。また、机上提示環境では、移動をしながら会議を行うわけではないため、話のきっかけを作るような参加者が必要であり、その結果、会議を進行する役となる参加者とそれ以外の参加者に分かれたことが考えられる。そのため、グループごとで参加者の間で発話時間に大きな差が生まれてしまったことが考えられる。このように、机上提示環境のような移動があまり行われないような環境では、立ち位置によっては見る/見られるといった対

面量には差が生まれ、会議を進行するために進行役のような参加者が表れ、発話時間にも差が生まれやすいのではないかと示唆された。

6. おわりに

本研究では、VR 会議における言語行動と見る/見られるといった対面量の対応関係を分析し、VR 会議では発話量によって対面量も変化するかを調査した。この調査を行うために、壁提示環境と机上提示環境の2つのVR環境を用意して比較実験を行った。実験で録画した一人称視点映像を用いて、参加者の見ていた/見られていた時間を集計した、また、会議の発話記録を書き起こし、発話時間を集計した。

分析では、見ていた時間と見られていた時間を散布図で可視化した。散布図のマーカの大きさは、それぞれの参加者の発話時間に対応させた。これにより、見ていた時間や見られていた時間によって発話時間はどのくらいの時間だったのかを理解することができた。また、会議環境ごとの傾向や、グループごとの傾向なども分析することができた。

壁提示環境では、参加者ごとで見ていた時間と見られていた時間ともに、大きな差はなくまとまった分布であることが分かった。また、グループごとで見ていた時間も見られていた時間ともにまとまった分布になることが分かった。さらに、発話時間には大きな差はなく、参加者たちは同程度の発話をする事ができていたことが分かった。これは、移動を行うことで、相手の様子を伺いながら会議に参加するため、見ていた時間と見られていた時間は同程度になったのではないかと考えられる。また、まとまった行動が多くなるため、発話機会も同程度あった可能性があることが示唆された。

机上提示環境では、参加者ごとに見ていた時間と見られていた時間ともに大きな差が見られ、ばらけた分布になることが分かった。さらに、参加者間で発話時間に大きな差が現れることも分かった。これは、参加者の立ち位置によって見ていた時間と見られていた時間に変化があることが原因であると考えられる。また、移動して会議を進めるわけではないため、会議を進めるきっかけとなる人が現れやすく、そのため進行をする参加者は発話時間が長くなってしまい、その他の参加者は発話時間が短くなってしまふことが考えられる。このように、机上提示環境では、立ち位置によって対面量には偏りが生じやすく、発話時間の差も生じやすくなる可能性があることが示唆された。

このように、壁提示環境では、参加者同士がまとまって行動しやすく、発話時間も均等になる可能性があることが考えられる。また、壁提示環境では見る/見られるといった対面量から発話がどの程度だったのかを予測できる可能性があることが示唆された。そのため、壁提示環境のよ

うな動きが促されたような環境では、対面量を見ることで会議により積極的に参加しているかが分かるのではないかと考えられ、会議を進める上での指標となり円滑に会議を進めることに繋がるのではないかと考えられる。また、VRで会議において壁提示環境のような環境では、全参加者が均等に発話しやすい環境である可能性が示唆されたため、会議環境を用意する上でどのような環境にデザインするかという際の手がかりになる可能性があるのではないかと考える。

一方、机上提示環境では、参加者間で見る/見られる時間は偏りが生じやすく、発話時間にも偏りが生じやすい可能性があることが示唆された。これは、参加者の立ち位置が関係している。発話時間が長いものにも関わらず見られていた時間が短い参加者がいた。この参加者は図6のような他の参加者の後ろ側から参加しており、このグループで最も発話時間が長く、他の参加者を見ながら会議の進行をしていた。このように、VR会議では後ろから話しかけ進行を進めるといったこともあるということが分かった。参加者の発話内容や会議の様子をより深掘りすることで、VR会議を進行する上での戦略をより理解することに繋がるのではないかと考えられる。



図6 発話時間が長く見られていた時間が短い参加者の立ち位置

今後は、会議に対する満足度と発話時間や対面量との関係を見ることを考えている。発話時間の差や対面量の差が生じる原因はその会議で納得していない面なども関係していると考えられる。そのためこれらの関係を調べることで、会議に対する満足度が言語行動や対面量に関係しているのか、それとも関係がないのかを知ることができると考える。この関係を調べることで、会議をより円滑に進め全員が納得できるような会議を行う手がかりを知ること繋がると思う。

謝辞 本研究は、北陸先端科学技術大学院大学の令和4年度萌芽的研究支援および科研費(21K11978、22H03634)の補助を受けた。

参考文献

[1] Mehrabian, A. and Ferris, S. R.: Inference of attitudes from nonverbal communication in two channels. *Journal of Consulting Psychology*. Vol.31, No.3, pp.248-252

(1967).
 [2] Kamada, K., Watarai, R., Wang, T. Y., Takashima, K., Sumi, Y. and Yuizono, T.: Explorative study of perceived social loafing in VR group discussion: A comparison between the poster presentation environment and the typical conference environment. *IFIP TC.13 International Conference on Human-Computer Interaction (INTERACT)*. LNCS. Vol.14144, pp.115-134 (2023). https://doi.org/10.1007/978-3-031-42286-7_7
 [3] Latané, B., Williams, K. and Harkins, S.: Many hands make light the work: The Causes and Consequences of Social Loafing. *Journal of Personality and Social Psychology*. Vol.37, No.6, pp.822-832 (1979).
 [4] 渡会隆哉, 鎌田光太郎, 角康之, 王子洋, 由井蘭隆也, 高島健太郎.: VR会議における他者との対面量に着目した会議の質評価の試み. *マルチメディア, 分散, 協調とモバイルシンポジウム 2023 論文集*. pp.1657-1665 (2023).
 [5] 榎本美香, 伝康晴.: 話し手の視線の向け先は次話者になるか. *社会言語科学*. Vol.14, No.1, pp.97-109 (2011).
 [6] Wallace, P. and Maryott, J.: The impact of avatar self-representation on collaboration in virtual worlds. *Innovate: Journal of Online Education*. Vol.5, No.5 (2009).
 [7] Han, E., Miller, M. R., DeVaux, C., Jun, H., Nowak, K. L., Hancock, J. T., Ram, N. and Bailenson, J. N.: People, places, and time: a large-scale, longitudinal study of transformed avatars and environmental context in group interaction in the metaverse. *Journal of Computer-Mediated Communication*. Vol.28, No.2 (2023).
 [8] 鎌田光太郎, 渡会隆哉, 王子洋, 高島健太郎, 角康之, 由井蘭隆也.: VRグループ・ディスカッションにおける会議形態の違いがターンテイキングに与える影響. *ヒューマンインタフェース学会研究報告集*. Vol.25, No.3, pp.97-102 (2023).
 [9] Lafferty, J. C. and Eady, P. M.: The desert survival problem. *Experimental Learning Methods*. Plymouth (1974).