

## GROW モデルに基づく目標達成にフォーカスしたコーチング対話システム

## GROW-Based Coaching Dialogue System Focused on Goal Achievement

山崎 紘介<sup>1)</sup> 櫻井 崇貴<sup>1)</sup> 長澤 史記<sup>1)</sup> 白松 俊<sup>1)</sup>  
Kosuke Yamazaki Takayoshi Sakurai Fuminori Nagasawa Shun Shiramatsu

## 1. はじめに

近年、個人のキャリア設計や地域での起業・活動支援の現場において、「やりたいこと」を言語化し、具体的な行動へとつなげる支援の重要性が高まっている。特に「さかさま不動産」[1]のように、借り手の「やりたいこと」を起点として物件の大家とのマッチングを行う取り組みでは、借り手自身がその想いや理想を整理し、借りるという行動へとつなげることができていないケースがあるという現場の声が寄せられていた。

他方で、雑談対話や傾聴対話といった領域では、大規模言語モデル (LLM) や AI エージェントを活用した対話システムの研究が進められており、そこではユーザの語りや自己開示を自然に引き出すことが重視されている[2][3][4]。その背景には、特定のタスクに縛られない柔軟な対話を得意とする LLM の特性[5]があり、自由なやりとりの中でユーザの深い内省を促す研究も行われている[6]。しかし、ユーザを特定の思考プロセスに沿って導いたり、目標設定や行動促進といった支援を行ったりするには、LLM の自由度の高さがかえって障壁となる可能性も指摘されている[7]。本研究に先立って構築した LLM による対話システムにおいても、ユーザが「やりたいこと」を語る中で話題が枝葉に広がりすぎてしまい、結果として「何をすればよいか分からない」とユーザに感じさせる場面がしばしば見られた。こうした行き詰まりを解消するために、GROW モデルによる対話の構造化が有効ではないかと考えた。

GROW モデルとは、ジョン・ウィットモアによって考案されたコーチング手法[8]であり、目標設定 (Goal) と現状分析 (Reality)、選択枝の探索 (Options)、意志決定 (Will) の 4 段階を通して、対話者の内省と行動変容を促すものである。本研究では GROW モデル (Goal, Reality, Options, Will) を対話の枠組みとして導入することで、対話が枝葉に広がることで具体的な行動の明確化まで進まない課題に対して、目標と現状の明確化を促し、効率的な目標達成支援を試みる。これにより、LLM による自由な対話の力を活かしつつも、GROW モデルによってその対話を構造化・焦点化することで、より個別的かつ実践的な支援が可能になると我々は考えた。

さらに GROW の各フェーズで引き出された情報は、ユーザの目標・課題・選択・意思といった背景情報に対応しており、LLM 対話システムにおける「個別化」の手がかりになると考えられる。対話 AI における個別化の精度向上にはユーザに関する事前情報の活用が重要であるという指摘[9]もあり、GROW に基づく対話ログは、そのような文脈においても有用な資源となる可能性がある。

本研究では、GROW モデルを組み込んだプロンプトを用いた LLM 対話システムを設計・実装し、研究室内で主観

的評価実験を行った。本論文ではその概要と初期的な検討結果を報告し、やりたいことの明確化支援における GROW モデルの有効性と、個別化対話の可能性について考察を行う。

## 2. 関連研究

目標達成支援を目的とした対話システムの分野において、コーチングの枠組みを取り入れる試みは近年注目を集めている。ビジネス領域では、LLM を AI コーチとして活用する事例が見られる[10][11]が、そうしたシステムの効果については十分に検証されていない。

一方、学術分野では、下田らが ELIZA[12]を基盤とした非タスク指向型対話システムに GROW モデルを導入し、ユーザの思考整理や行動支援を目的とした対話構造を提案した[13]。彼らの手法では、固有表現抽出や極性判定などの自然言語処理技術を活用し、テンプレート応答とフレームベースの対話管理によって、GROW の 4 フェーズ (Goal, Reality, Options, Will) に対応した対話制御を実現している。加えて、Li らの研究では、LLM が確率モデルによる言語生成を行っているにすぎないことから、LLM を用いたタスク達成型の対話システムにおいて、LLM にサブゴール、すなわちゴールに至るまでの過程を明確化させることが重要であると述べている[7]。また、杉田らは、運転行動の学習支援ロボットに GROW モデルを応用し、危険状況の認識とその対応方法の記憶定着を支援するシステムを提案した[14]。本研究は、下田らの構造的対話設計アプローチを踏襲しつつ、LLM の持つ自由度と表現力を活用した点に特徴がある。こうした構造化手法の違いは、LLM の柔軟性を活かしつつ、対話に方向性を与える枠組みとして GROW を活用することの実装的な価値と、そこに潜む検証の意義を示唆するものである。加えて、ユーザの目標達成に向けて、GROW の枠組みに基づいた構造化が対話のタスク達成に寄与する可能性を検証する。また、ロボットによる支援とは異なり、対話システムを基盤とした GROW 実装は、より汎用的かつ展開可能な支援の在り方を模索する新たなアプローチと位置づけられる。

## 3. 提案手法

本研究では、ユーザが自身の「やりたいこと」や目標を明確化・深掘りできるよう支援する対話システムを目指した。具体的には、目標達成の支援手法として広く知られる GROW モデル (Goal, Reality, Options, Will) を、大規模言語モデル GPT-4o にプロンプトとして組み込むことで、構造的かつ柔軟なコーチング対話を実現することを試みた。

## 3.1 システム構成

システムは、ユーザが Web ブラウザからアクセスできる自作の UI (図 1) を通じてテキストベースの対話が行える構成とし、OpenAI API を用いて GPT-4o と連携する形式

1) 名古屋工業大学 Nagoya Institute of Technology.

で実装した。ユーザはチャット上でやり取りを行い、対話の履歴が逐次記録・取得される。

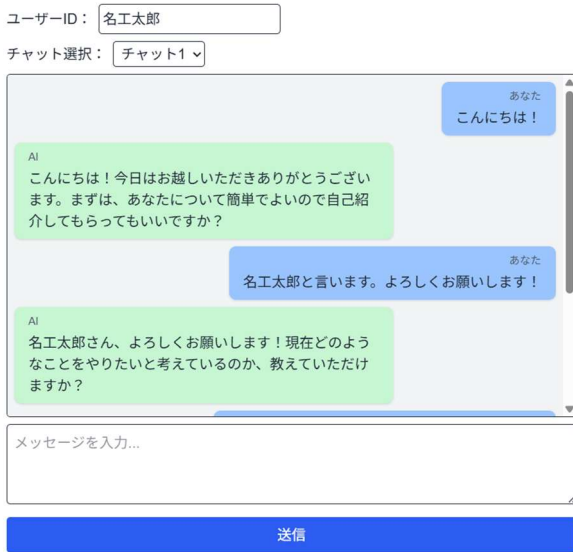


図1 ユーザインタフェースの画面 (GROWあり対話UIの画面例)

### 3.2 GROWモデルの導入とプロンプト設計

本研究では、GROWモデルの導入効果を検証するために、対話システムのLLMに与えられるプロンプトについて「GROWあり」と「GROWなし」の2種類のモードを用意した。「GROWあり」は、システムがGROWの各フェーズに沿って対話するモードであり、質問例を表1に示す。「GROWなし」は、構造的な指示を与えず、自由に聞き取りを行うものである。

両モードにおいて、AIには「インタビュアー」としての役割をプロンプトにより明示し、ユーザのやりたいことについて語りを引き出す対話を複数ターンで行うよう設計した。

なお、本プロンプトでは図2で示すように、GROWモデルの各フェーズは明示的にユーザに提示せず、自然な会話の流れの中に段階的に埋め込む形で構成された。

両モードにおいて、最終的には対話ログをもとにユーザ情報を要約するように設計した。要約フォーマットも

表1 GROWの各フェーズにおける質問例

GOAL (目標)	「未来のあなたはこの事業を通してどんなゴールを達成したいですか？」
REALITY (現状)	「これまでゴールに対して何をを行い、どのような成果を出しましたか？」
OPTIONS (選択肢)	「障害がなければ何をしますか？他にできることはありますか？」
WILL (意思)	「自分がやろうと決めた行動を最後までやり通す確率は、1から10で評価するといくつですか？」

GROWの有無に応じて異なり、GROWありモードでは各フェーズに沿った構造的な整理を行い、GROWなしモードでは自由記述的な形式で出力させた。

#### # 対話の進め方

##### 1. 以下の質問から始めて、ユーザの基本情報を得る

- 「まずは、あなたについて簡単でよいので自己紹介してもらってもいいですか？」
- 「お名前などあなたについてお聞きしてもいいですか？」

##### 2. GOAL(目標や理想)について聞く

- 「これからこの事業を通して達成したい目標はありますか？」

##### 3. Reality(客観的なスタート地点)を特定する

- 「今までのところ、ゴールに対して何を行い、どのような成果を達成しましたか？」

##### 4. Options(選択肢)について解決策を考える

- 「もしも、このゴールを達成することへの障害が全く何もないとしたら、あなたは何をしますか？」

##### 5. Will(意志)や、実行責任を果たす気持ちを確認し、解決策を具体的に明確な行動を促す

- 「自分がやろうと決めた行動を最後までやり通す確率は、1から10で評価するといくつですか？」

図2 GROWありのLLMに与えたプロンプト(一部抜粋)

### 4. 実験の設計と実施

本研究では、GROWモデルに基づく対話システムの構造化が、ユーザの目標明確化や個別性の知覚に与える影響を探索的に検討するため、GROWありモードを導入した対話(チャット1)と、GROWなしモードを導入した対話(チャット2)を比較する実験を行った。

#### 4.1 実験概要

被験者は研究室メンバー12名であり、各参加者が「夢ややりたいこと」として2つのテーマを挙げ、それぞれに対してチャット1とチャット2の2種類を体験した。対話の順番はランダムに割り当てられ、順序効果を排除した設計とした。

#### 4.2 条件設定と対話構成

本実験では、GROWモデルなしの条件(チャット2)をベースラインとし、GROWモデルを導入した対話(チャット1)がどのような違いを生み出すかを検証した。両条件とも、使用したモデルはgpt-4o-2024-08-06であり、実験は3.1節で述べた自作のWeb UI環境上で実施された。評価項目は、24項目あり、表2において、内訳を示す。これらの項目は下田らの先行研究と比較可能なユーザ体験指標(Q1~Q9)と、本研究独自に設定したGROWの各フェーズや個別化支援・伴走支援の感覚に対応する指標(Q10~Q24)で構成されている。前者には比較のため5段階のリッカート尺度、後者には7段階のリッカート尺度を用いた。なお、Q20、Q22では被験者の回答バイアスを軽減するため、逆評価を用いた。また、各項目についてはスチューデントのt検定で有意差を調べた。

表 2 設問項目の一覧

ユーザ 体験 (先行研究 比較)	Q1	システムの応答には意味がある
	Q2	対話が楽しかった
	Q3	自然な応答だった
	Q4	システムに温かみを感じた
	Q5	もう一度話したいと思った
	Q6	対話が弾んでいると感じた
	Q7	応答が簡潔で分かりやすかった
	Q8	システムが話を理解してくれていると感じた
	Q9	自分の考えが整理されたと感じた
GROW フェーズ	Q10	目標が明確になった
	Q11	自身の現状について理解が深まった
	Q12	目標の実現に向けての選択肢が広がった
	Q13	やるべき行動や意志が明確になった
伴走感 信頼感	Q14	この対話相手は「一緒に考えてくれている」と感じた
	Q15	この対話相手は「自分の気持ちや考えに寄り添ってくれている」と感じた
	Q16	この対話相手は「信頼できる対話パートナー」だと感じた
対話進行	Q17	この対話は「目標→現状→選択肢→行動」という流れがわかりやすかった
	Q18	対話の進行が迷わずに進められる流れだったと感じた
個別化	Q19	この対話相手の問いかけや提案は、あなたの状況や話に合わせてくれていると感じた
	Q20	この対話が誰にでも同じように話しているだけだと感じた
自由度	Q21	対話中、自由に話したり、話題を広げられる余地があったと感じた
	Q22	対話が「決められた質問に答えるだけ」のように感じる瞬間があった
行動意欲	Q23	対話を通じて、次にやるべきことや行動の方向性が見えたと感じた
	Q24	対話を終えて、実際に何か一歩動いてみようと思った

### 4.3 実験と検証

#### 4.3.1 先行研究との比較

下田らは、ELIZA ベースの非タスク指向型対話システムに GROW モデルを導入し、ユーザの思考整理や行動支援を目的とした対話設計を行った。彼らの研究では、対話の「温かさ」や「弾み」、「再対話意欲」をはじめとする 9 つの主観的評価項目において、GROW モデル導入による対話支援の有効性の傾向が報告されている。本研究では、下田らが用いた評価項目に対応する主観質問項目 (Q1~Q9) を設定し、GROW モデルに基づく対話構造 (チャット 1)

と自由形式の対話 (チャット 2) の比較を行った。これにより、先行研究と同様の効果が LLM ベースの対話でも再現されるかを検証した。

#### 4.3.2 GROW モデルの各要素の実装

GROW モデルの各フェーズ (Goal, Reality, Options, Will) が対話内でどのように機能したかを検証するため、Q10~Q13 の評価項目を用いてチャット 1 (GROW あり) とチャット 2 (GROW なし) の比較を行った。これらの項目は、各フェーズに対応する問いかけがユーザの目標設定や自己理解、選択肢の探索から意思決定までにどのような影響を与えるかを測る目的で設定されている。

#### 4.3.3 コーチング体験の実現

GROW モデルの有無がユーザにどのようなコーチング体験を与えたかを把握するため、Q14~Q24 の主観評価項目を用いて調査を行った。これらの項目は、伴走感や信頼感、対話の流れのわかりやすさ、個別対応の印象、行動意欲の変化など、多面的な観点から対話体験を評価するものであり、特にコーチングにおいて重視される共に考える姿勢や主体的な次の一歩の形成に深く関わる。

## 5. 実験結果

4.3 節での検証項目に基づき、先行研究で言及されていたユーザ体験 (Q1~Q9)、GROW の実装 (Q10~Q13)、コーチング体験 (Q14~Q24) の 3 つに設問を分類した結果をまとめた。

### 5.1 ユーザ体験における評価結果

ユーザ体験に関する 9 項目の主観評価の結果を表 3 に示す。評価尺度に関しては、Q7「応答が簡潔で分かりやすかった」を除くすべての項目において、チャット 1 の平均スコアがチャット 2 を上回った。

具体的には、Q1「システムの応答には意味がある」でチャット 1 は 4.00、チャット 2 は 3.42 といった差が見られた。唯一、Q7 においてはチャット 2 (3.67) がチャット 1 (3.50) を上回る結果となった。

スチューデントの  $t$  検定の結果、有意差が確認されたのは Q1 ( $p = 0.0271$ ) のみであり、その他の設問では有意差は確認されなかった。ただし、Q5「もう一度話したいと思った」 ( $p = 0.0558$ ) および Q9「考えが整理されたと感じた」 ( $p = 0.0501$ ) については、有意水準に近い値が示された。総じて、ユーザ体験において、一部設問に有意差があったが、先行研究で示された明確な差はなかった。

表 3 ユーザ体験における評価(5段階評価)

設問 番号	チャット 1 尺度平均	チャット 2 尺度平均	$p$ 値
Q1*	4.00	3.41	<b>0.0271</b>
Q2	2.83	2.50	0.3125
Q3	3.83	3.58	0.5152
Q4	3.66	3.41	0.4627
Q5	3.17	2.41	0.0558
Q6	3.25	2.67	0.1106
Q7	3.50	3.67	0.8125
Q8	4.33	3.83	0.1484
Q9	4.08	3.17	0.0501

\* $p < 0.05$

## 5.2 GROW モデルの各要素に関する評価結果

GROW モデルの各フェーズ (Goal, Reality, Options, Will) に対応する主観評価項目の結果を表 4 に示す。全ての項目において、GROW モデルを導入したチャット 1 の平均スコアは、導入していないチャット 2 の平均スコアを上回った。具体的には、目標の明確化に関する Q10 (チャット 1: 5.50, チャット 2: 4.50), 現状理解に関する Q11 (5.25, 4.50), 選択肢の広がりに関する Q12 (4.42, 3.42), および行動や意志の明確化に関する Q13 (5.42, 4.00) において、いずれもチャット 1 の方が高い評価を得た。しかし、スチューデントの  $t$  検定を適用した結果、これら 4 項目のいずれにおいても統計的に有意な差は確認されなかった。  $p$  値はいずれも 0.05 を上回っており、GROW 導入によるフェーズごとの評価向上は確認されたものの、有意差としての判断には至らなかった。

表 4 各 GROW フェーズに関する評価(7段階評価)

設問番号	チャット 1	チャット 2	$p$ 値
	尺度平均	尺度平均	
Q10	5.50	4.50	0.1530
Q11	5.25	4.50	0.1574
Q12	4.42	3.42	0.0892
Q13	5.42	4.00	0.0978

## 5.3 コーチング対話の評価結果

コーチング体験に関する 11 項目の主観評価を表 5 に示す。これらの項目において、肯定的な評価が高得点で示される通常 9 項目では、すべての設問においてチャット 1 (GROW あり) の平均スコアがチャット 2 (GROW なし) を上回った。

また、否定的な印象を測る逆評価項目 (Q20「誰にでも同じように話しているだけだと感じた」、Q22「決められた質問に答えるだけのよう感じた」) においては、スコアが低い方が好ましいとされるが、いずれもチャット 1 がチャット 2 より低いスコアを示した。このことから、全 11 項目においてチャット 1 がより良好な評価を得た。

スチューデントの  $t$  検定の結果、行動形成に関する設問

表 5 コーチング体験における評価(7段階)

設問番号	チャット 1	チャット 2	$p$ 値
	尺度平均	尺度平均	
Q14*	5.08	3.75	<b>0.0465</b>
Q15	4.83	4.50	0.6013
Q16	4.67	4.00	0.3388
Q17	5.33	3.67	0.0512
Q18	5.33	4.67	0.1513
Q19	4.83	4.42	0.4102
Q20(逆評価)	3.42	4.42	0.1184
Q21	4.58	4.33	0.6230
Q22(逆評価)	3.92	4.25	0.5695
Q23**	5.91	3.83	<b>0.0064</b>
Q24*	5.75	4.08	<b>0.0341</b>

\* $p < 0.05$  \*\* $p < 0.01$

Q23「次にやるべきことや行動の方向性が見えた」 ( $p = 0.0064$ ), Q24「実際に何か一歩動いてみようと思った」 ( $p = 0.0341$ ), および伴走感についての設問 Q14「一緒に考えてくれていると感じた」 ( $p = 0.0465$ ) の 3 項目において、チャット 1 が有意に高い評価を得た。

一方で、個別対応や対話の自由度に関連する設問では、いずれもチャット 1 がチャット 2 を上回る傾向を示したものの、統計的な有意差は確認されなかった。

## 6. 考察

### 6.1 先行研究との比較と考察

下田らの先行研究では、9 つの主観的評価項目において、GROW モデルを導入した ELIZA の対話システムの有効性が報告されている。本実験では、Q1「応答に意味があると感じた」で有意に高い評価が得られた。これは、GROW モデルに基づいた構造的な問いかけが、ユーザに対して「対話を通じて考える」という感覚を強化した可能性を示している。このことは、単に応答が自然であるだけでなく、「対話に目的や価値がある」と感じられるためには、一定の枠組みが必要であることを示唆する。

一方で、その他の項目では明確な差は出なかったが、下田らの GROW なし条件 (ELIZA ベース) では、図 3 で示すように、主観評価項目が平均 1 点台後半にとどまっていたことが報告されていた。本研究の GROW なし条件 (チャット 2) がそれに比して平均 3 点前後を示したことは、LLM が持つ自然な応答能力が、従来のテンプレート型より高いベースラインを形成している可能性がある。結果として、GROW の効果が顕著に現れにくい可能性もあり、意味を与える構造が LLM にどのように重なり得るかという観点が今後の検討課題である。

また、Q7「応答が簡潔で分かりやすかった」においては、チャット 2 がチャット 1 を上回った。これに関して、アンケートの自由記述では「1 発話に複数の質問が含まれていた」「応答が長く読みづらい」といった指摘がみられた。このことから、GROW モデルに基づく構造的対話では、各フェーズを網羅しようとする意図がかえって 1 ターンあたりの情報量を増やし、結果として応答の過剰説明や質問の過密化を招いた可能性がある。応答の複雑化は、ユーザにとって処理負荷となり、対話の簡潔さを損なう一因となり得る。一方で、GROW なし条件では自由対話ゆえに問いかけが単発で済む場面が多く、結果としてシンプルかつ理解

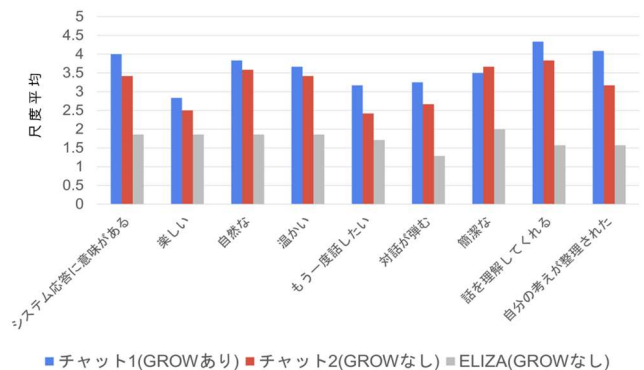


図 3 先行研究[13]ベースラインとのユーザ評価比較

しやすい応答が実現されたと推察される。この結果は、GROW 導入において「問いの構造化」と「応答の簡潔さ」とのバランスをいかに取るかが今後の課題であることを示唆している。たとえば、1 ターンにおける問いの数を制限する設計や、各フェーズの進行を明示的に複数ターンに分割するプロンプト構造を取り入れることで、ユーザの負担を軽減しながら構造的支援を実現する工夫が求められる。

## 6.2 GROW モデルの各要素に関する考察

GROW モデルの各フェーズ (Goal, Reality, Options, Will) が対話内でどのように機能していたかを検討するため、Q10～Q13 の主観評価項目に対して統計的検定 (スチューデントの  $t$  検定) を実施した。結果として、すべての項目において GROW あり条件 (チャット 1) のスコアが GROW なし条件 (チャット 2) を上回ったが、有意差が確認された項目はなかった。これらの結果は、GROW の導入においてユーザ側で意識的に GROW を意識した対話が可能でなかった可能性がある。

ただし、Q12「目標の実現に向けての選択肢が広がった」( $p = 0.0892$ ) や Q13「やるべき行動や意志が明確になった」( $p = 0.0978$ ) といった後半フェーズ (Options・Will) に対応する項目では、GROW あり条件 (チャット 1) のスコアが 1 点以上 GROW なし条件 (チャット 2) を上回っていた。このような評価の差は、GROW の後半フェーズにおける構造的な問いかけが、行動を意識させる対話展開につながっていた可能性を示す一つの観察結果であり、今後の検討に向けた示唆といえる。

総じて、GROW モデルの導入は各フェーズでスコアは全体として GROW あり条件の方が高かったが、統計的に有意な支援効果を示すには至らなかった。ただし、対話全体の進行構造に関する Q17「目標→現状→選択肢→行動という流れがわかりやすかった」において、チャット 1 の方が高いスコアを示した ( $p = 0.0512$ )。これは、GROW に基づく対話設計が一定の構造的ガイドとして機能していた可能性がある。こうした構造的問いかけの設計が、LLM ベースの自由対話に一定の方向性と整理をもたらす可能性は示唆された。本対話システムでは、GROW 構造を対話の中に自然に埋め込む形式を採用し、各フェーズをユーザに明示的に提示することは行わなかったが、対話全体の進行としては GROW モデルに沿った流れが一部ユーザに認識されていた可能性もある。今後の検討として、各フェーズの構造化をユーザが明確に認識できる粒度で提示するようなプロンプト設計を導入することで、GROW の各フェーズにおける支援効果がより明瞭となり、対話を通じたコーチングの再現性を高めることが期待される。

## 6.3 対話の伴走感と支援感の考察

主観評価の分析において、行動意欲や伴走感に関する一部の項目で統計的に有意な差が確認された。GROW モデルによる対話設計は、一定の「伴走感」や「行動への動機づけ」を与える上で有効に機能した可能性がある。

特に、Q23「次にやるべきことや行動の方向性が見えた」および Q24「実際に何か一歩動いてみようと思った」といった行動意欲に関する項目において、GROW あり条件 (チャット 1) が有意に高い評価を得たことは、構造的な問いかけがユーザの内省を促し、自己決定や目標達成への行動

を支援するコーチング的機能を果たしていたことを示唆している。

また、Q14「一緒に考えてくれている」と感じたという伴走感に関する項目でも有意差が見られた点から、GROW モデルの導入が本対話システムを「目標達成に向けて共に歩むコーチ」的な存在としてユーザに認識させることに一定の寄与をしていたと考えられる。

このような結果は、本研究が志向する「GROW モデルに基づく目標達成にフォーカスしたコーチング対話システム」において、行動の方向づけだけでなく、ユーザに寄り添いながら伴走する支援の在り方が一定程度実現されていたことを示唆するものである。

## 6.4 対話の個別化の考察

「個別化」に関する評価について、明確な効果を示す統計的な結果は得られなかった。Q19「問いかけや提案が自分の状況に合っていた」では、GROW あり条件 (チャット 1) の平均スコアが GROW なし条件 (チャット 2) を上回ったものの、有意差は確認されなかった ( $p = 0.3924$ )。また、逆評価項目である Q20「誰にでも同じように話しているだけだと感じた」においても、チャット 1 の方がやや低いスコアを示すにとどまり、統計的な差異は見られなかった。

この結果は、ユーザごとの文脈や状況に応じた柔軟な対応、すなわち「個別化された対話」の実現という観点において、現段階の設計では十分な対応がなされていたとはいえないことを示している。特に、本研究では GROW の 4 フェーズをユーザに明示的に提示することなく、プロンプトの内部構造として対話に自然に埋め込む形式を採用した。これは、ユーザが GROW 構造を意識することなく、目標や現状、選択肢、意志といった語りを自然に引き出すことを意図した設計である。しかしながら、取得された情報をもとにユーザごとの背景や文脈に踏み込んだ問いかけを行うには、LLM の応答の粒度や問いかけの最適化において十分とは言えず、結果として「個別化されている」という実感にはつながりにくかった可能性がある。

ただし、本研究を通じて得られた GROW の各フェーズに関する情報 (目標、現状、選択肢、意志) は、今後の対話設計において個別性を高める文脈情報として活用できる可能性がある。これらの情報をもとに、ユーザの発話よりの確に回答する設計や、共感的な応答生成を工夫することで、「一人ひとりに寄り添う」感覚の強化につながれると考えられる。

加えて、Q21「自分のペースで自由に話すことができた」に関する評価では、チャット 1 (GROW あり) の平均スコアがチャット 2 を上回っており、統計的な有意差は見られなかったものの、GROW による対話構造が自由な語りを妨げるものではなかったことが示唆される。これは、プロンプト内に GROW の問いかけを埋め込むという設計によって、構造的支援とユーザ主導の語りが両立可能である可能性がある。この結果は、今後より明示的に GROW のフェーズを提示したとしても、自由度の体験が損なわれるとは限らない可能性を示しており、GROW の構造化粒度を調整する方向でのさらなる設計改善に向けた足がかりともなりうる。総じて、GROW モデルの導入は「個別化された体験」を統計的に明示するには至らなかったものの、この結果は、

非構造的な LLM 対話に構造的フレームを取り入れることで、より深い文脈理解や個別対応への足がかりが得られる可能性を示すものである。

## 7. おわりに

本研究では、大規模言語モデル (LLM) を用いた対話システムに、コーチング手法である GROW モデルを導入し、その構造的対話がユーザの目標形成や行動意欲、対話の個別化支援にどのような影響を与えるかを探索的に検討した。

分析の結果、GROW ありの条件 (チャット 1) では、「次にやるべきことが見えた」「実際に行動してみようと思った」といった行動意欲に関する項目において有意差が見られた。また Q17 の GROW (Goal, Reality, Options, Will) の流れに関する評価では、 $p$  値が 0.05 に近い水準 ( $p = 0.0512$ ) を示しており、構造的対話の理解がある程度反映されていた可能性がある。これにより、LLM 対話に GROW モデルの構造的問いかけを導入することが、ユーザの行動を後押しする一助となる可能性が示唆された。一方で、GROW の各フェーズに対応する評価項目では、全体としてスコアの上昇傾向は見られたものの、有意差は確認されなかった。特に初期フェーズ (Goal・Reality) では構造化の効果が数値上に明瞭に表れにくく、プロンプト設計における構造化粒度や誘導の工夫に課題が残ることが示唆された。さらに、対話の個別化に関する実感については、有意差は見られなかったが、GROW 構造によって得られたユーザの文脈情報 (目標、現状、選択肢、意志) をもとに、より柔軟かつ最適化された問いかけや共感的応答を設計する余地が確認された。また、GROW の問いを通じて引き出された情報が、継続的にコーチングを行う上で対話の個別性や伴走感を支える基盤となりうる可能性はある。なお、本研究の実験参加者は研究室のメンバーに限定されており、ユーザ層に一定の偏りがあったことは限界の一つである。今後は、たとえば「さかさま不動産」の実際の利用者など、より多様な属性や背景を持つユーザを対象とした実証研究を行うことで、本手法の汎用性や実用性をより広く検証していく必要がある。また、本研究で収集した自由記述の内容や対話ログを構造的に分析することで、GROW モデルに基づく支援がユーザの語りにどのような影響を与えたのか、また、どのような設計上の課題が残されているのかをより精緻に捉えることが可能となる。

今後は、GROW モデルに基づいて得られるユーザ情報 (目標、現状、選択肢、意志) を、ユーザの背景理解やニーズ把握に活用することが、個別最適な対話支援の実現につながるという仮説をもとに、さらなる設計の工夫を重ねていきたい。たとえば、個別化されたプロンプトの動的生成や、対話履歴を踏まえた継続的な対話設計、AI による「伴走的支援」の信頼性向上といった方向性が今後の課題として挙げられる。LLM に GROW のような構造的枠組みを組み込むことで、単なる雑談や一回限りの応答にとどまらず、ユーザとともに目標に向かって歩むような目的志向型のコーチングの対話の可能性が拓かれることが期待される。

## 謝辞

本研究は、株式会社 On-Co, JST CREST (JPMJCR20D1), JSPS 科研費 (24K03052, 25K00866), 篠原欣子記念財団の支援を受けたものです。さかさま不動産の現場での知見から貴重なご助言を頂いた On-Co の皆様に深謝します。

## 参考文献

- [1] 株式会社 On-Co : さかさま不動産, さかさま不動産 (オンライン), 入手先 (<https://sakasama-fudosan.com/>) (参照 2025-06-05) .
- [2] Lee, Y. C., Yamashita, N., Huang, Y., & Fu, W. (2020, April). "I hear you, I feel you": encouraging deep self-disclosure through a chatbot. In *Proceedings of the 2020 CHI conference on human factors in computing systems* (pp. 1-12).
- [3] 鈴木香保, 上乃聖, & 李晃伸. (2024). カウンセリングにおける悩み構造の言語化を支援する音声対話システム, HAI シンポジウム 2024 ポスター発表 P-20.
- [4] 于, 海博, 白松, 俊, 橋本, 慧海, & 中野, 幹生. (2025). 困り事を言語化できない子ども向けの傾聴対話システムの開発. 信学技報, 124(447), 48-53. <https://www.ieice.org/iss/Consen2024-15> [Consen2024-15]
- [5] Yi, Z., Ouyang, J., Liu, Y., Liao, T., Xu, Z., & Shen, Y. (2024). A survey on recent advances in llm-based multi-turn dialogue systems. arXiv preprint arXiv:2402.18013.
- [6] Park, S., & Kulkarni, C. (2023). Thinking assistants: Llm-based conversational assistants that help users think by asking rather than answering. arXiv preprint arXiv:2312.06024.
- [7] Li, H., Leung, J., & Shen, Z. (2024). Towards Goal-oriented Prompt Engineering for Large Language Models: A Survey. arXiv preprint arXiv:2401.14043.
- [8] Whitmore, J. (1996). \*Coaching for performance\*. Nicholas Brealey Publishing.
- [9] Purificato, E., Boratto, L., & De Luca, E. W. (2024). User modeling and user profiling: A comprehensive survey. arXiv preprint arXiv:2402.09660.
- [10] COACHA : AI コーチングサービス, COACHA (オンライン), 入手先 (<https://www.coacha.com/services/ai.html>) (参照 2025-06-05) .
- [11] Forbes JAPAN 編集部 : AI コーチングが人材育成の常識を変える, Forbes JAPAN (オンライン), 入手先 (<https://forbesjapan.com/articles/detail/75952>) (参照 2025-06-05) .
- [12] Weizenbaum, J. (1966). ELIZA—a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, 9(1), 36-45.
- [13] 下田耕太郎, & 西野哲朗. (2019, November). コーチングを行う対話システムの実現に関する研究. In 人工知能学会研究会資料 言語・音声理解と対話処理研究会 87 回 (2019/12) (p. 27). 一般社団法人人工知能学会.
- [14] 杉田渉, 加納政芳, 早瀬光浩, & 吉川大弘. (2024). GROW モデルを基に正しい運転行動を教えるロボットの基礎検討. In *人工知能学会全国大会論文集 第 38 回 (2024)* (pp. 3Xin214-3Xin214). 一般社団法人人工知能学会.