

大規模言語モデルを活用した 認知再構成法支援チャットボットシステム

梅田 悠哉¹ 楊 期蘭¹ 平野 真理² 矢谷 浩司¹

概要: 認知再構成法は、人々の精神状態を改善するのに有効な、確立された心理学的技法である。ユーザがチャットボットから助言を受けることや、逆にユーザがチャットボットに対して支援を行うことは、どちらもユーザの認知再構成を支援するのに効果的であることが先行研究によって示されている。しかしながら、これらのチャットボットとのやり取りが、ユーザの認知再構成や心理状態に対してもたらす影響の違いについての比較検証は、未だ十分になされていない。この課題に取り組むため、本研究では大規模言語モデルを活用して2種類のチャットボットを開発し、それらを組み込んだウェブシステムを構築した。一つは、ユーザと同様の悩みをユーザに相談するピアチャットボットであり、もう一つは、ユーザの悩みに助言を行うコンサルティングチャットボットである。統制群・ピア群・コンサルティング群の3群に分けられた計102名の参加者に対し、3週間にわたってユーザ実験を実施した。結果として、先行研究とは異なり、いずれの群においても参加者の心理状態に有意な変化は見られなかった。また、ピア群では、チャットボットの介入によって悩みに対する別の考え方を記述することがより難しく感じられたことが確認された。一方コンサルティング群においては、チャットボットの介入によって主観的解決に至った悩みの個数が有意に増加した。先行研究と照らし合わせながら、これらの結果が得られた理由について考察する。また、その考察を踏まえ、ユーザの認知再構成を効果的に支援するチャットボットシステムの設計に関しても議論する。

A Chatbot System for Supporting Cognitive Reframing Using a Large Language Model

YUYA UMEDA¹ CHI-LAN YANG¹ MARI HIRANO² KOJI YATANI¹

1. 序論

うつ症状や不安の軽減、心理状態の改善に有効な心理技法として認知再構成法がある [5]。これは、自身の抱えている悩みについて、状況を別の視点から捉えなおし、新しい建設的な考え方を探ることで、思考の柔軟性を高めるものである。例えば、「上司とすれ違ったが挨拶をされなかった」という状況について、「なぜ彼は怒っているのだろうか」という思考に囚われ、不安感を抱いてしまう場合があるだろう。このとき、状況をもう一度見直して、「もしかすると、忙しくて私に気づかなかっただけかもしれない」というような、より心理的負担の少ない捉え方もできる。こ

のように、別の視点からも状況を考える力を身につけることが、より建設的な考えを発見し、心理状態を向上させることに繋がる。

自身の考えから離れ別の視点を発見する、という作業は、一人で行うには困難が伴う [20] が、他人と会話を行うことで、人々は自身のこれまでの視点からは離れた別の視点を獲得することができる [3]。悩み相談においては、例えば助言を受ける側として、人々は他者から直接別の考え方を学ぶことができる [10]。また、逆に助言を与える側であっても、人々は他者への助言の提供を通じて客観的に別の視点を発見することができる [15, 22]。このように、人々は他者との会話を通じてネガティブな思考を再構成することができる。しかし、一部の人は、そもそも悩み相談をできる相手が周りにいなかったり、また悩みという非常に個人

¹ 東京大学 Interactive Intelligent Systems Laboratory

² お茶の水女子大学

的な問題を、他者に相談することに抵抗感を抱くことがある [14].

チャットボットはこのような課題に対する一つの解決策である。ユーザのメンタルヘルスケアを支援するチャットボットはこれまでに様々開発されており [12,21,28], これらのシステムは会話を通じてユーザに対し有益な情報や助言・指示を提供する。さらに、先行研究 [19] では、助言を「受ける」側としてのソーシャルボットの効果が検証され、ユーザがボットに対し助言を与える場合であっても、間接的にユーザの認知再構成を支援することができることが示された。しかしながら、それらのシステムの実際的な実現可能性であったり、異なる種類のやり取りがユーザに与える影響の違いについての理解は、未だ十分に明らかになっていない。

そこで本研究では、次の2つの問いを設定した。

RQ1.異なる種類のチャットボットとのやり取りが、自身の悩みに対するユーザの認知再構成にどのような影響を与えるか。

RQ2.異なる種類のチャットボットとのやり取りが、ユーザの心理状態にどのような影響を与えるか。

これら2つの問いを検証するため、102人の参加者を対象に22日間のユーザ実験を実施した。ユーザ実験では、各参加者を「チャットボットなし」の統制群、「ユーザがチャットボットに支援を行う」ピア群、「ユーザがチャットボットから支援を受ける」コンサルティング群の3つに分け、各群の結果を比較した。そして、各介入が参加者の心理状態と認知再構成の質に与える影響の違いを検討した。

結果としては、先行研究 [19] のような有意な心理状態の改善はどの群でも見られなかった。また、ピア群では、予想に反し、チャットボットの介入によって悩みに対する別の考え方を書くことが有意に難しく感じられたことが確認された。一方、コンサルティング群では、チャットボットの介入によって、参加者が解決したと感じた悩みの個数が有意に増加した。

本論文では、ユーザ実験と先行研究からの知見に基づき、これらの結果が得られた原因について考察する。その後、認知再構成を支援するチャットボットシステムを構築するにあたって、重要と考えられる要素について議論する。

2. 関連研究

2.1 認知再構成法

精神疾患や日々のストレスを改善する手法として、CBT (認知行動療法) が存在する。CBT はうつ病や不安症状を抱える患者に対して効果的であることが示されており [16], 心理療法の代表的な手法として広く用いられている [13].

「認知再構成法」は、CBT の主軸の一つである [9]. 認知再構成法では、自身が抱えるネガティブな思考から離れ、状況に対する別の見方を見つけ出し、考え方の柔軟性を高

めるという作業が行われる。これは、人々の悩みやストレス、不安、うつ症状を軽減するための確立された心理技法である [5]. 認知再構成法に熟達することは、心理状態を改善する上で重要であることが示されている [23].

しかしながら、自身の抱く思考から離れ、他の視点を見つめるという作業は、一人で行うには困難が伴う [20]. 特に、ネガティブな思考に悩まされている場合、その作業には一層の苦勞を要する場合がある。それは、ネガティブな思考が自己内省を促進し、自分自身へ向ける注意を強化する [27] ことによる。このような注意は、時には理想と現実の差を埋めようとする動機となる [7]. しかしながら、それが過度になるとネガティブな感情が強化され、それによってさらに自己への注意が強化される、という悪循環が形成されてしまい、持続的なうつ症状につながることもある [26]. このような内部に注意が向いた悪循環を解消する上では、他者との悩み相談といった外部からの支援が重要な役割を果たすと考えられる。

2.2 認知再構成における会話の利点

会話を通じた意見交換・共有によって、人々は多様な視点を獲得できる [3]. 例えば、悩み相談は会話の一形態である。悩み事を抱えているときには、アドバイザーと話すことで安心感を得ることができる [10]. アドバイザーとの会話を通じて異なる視点を学ぶことで、自身の問題を再評価できるようになるためである。

また、悩みを抱えている人と話すことは、支援を行う人にとっても有益であることが示されている。他者を助けることによって、支援を行った人自身のうつ症状が減少することが明らかになっていて [11], 神経学的にも他者を支援することはストレスに関係する脳内活動を減少させることが示されている [17]. また一般的に、人々は他人に対しては合理的な助言を行うことができる場合が多い。この現象は「ソロモンのパラドックス」と呼ばれている [15]. この現象が生じる理由は、問題が他人のものである場合、人は問題に対して心理的距離を保つことができ、冷静かつ客観的に状況を見ることができからである [22]. したがって、自分の抱く悩みと同様の悩みを語る他人に対して助言を行う場合、自分自身の悩みに対しても、冷静かつ客観的な視点から考えることができるようになる可能性がある [19].

しかし、悩み相談を行う、あるいは悩み相談に乗るという行為を日常生活において実践するには課題もある。一つは、悩み相談の相手が周りにいない人もいるということである。また、自身の失敗や欠点を他者に明かすことによって自身の評判や印象を損なうことを恐れ、周りの人に悩み相談を行うことに対して抵抗感を抱く場合もある [14].

2.3 認知再構成を支援するシステム

先行研究では、ユーザの認知再構成を支援する手法が

さまざまな提案されている。Smith ら [29] は、Flip*Dubt と呼ばれるウェブプラットフォームを構築した。これは、ユーザがお互いにお互いの問題に対してより建設的な考え方を提供するシステムである。しかし、このシステムでは人手によって問題に対する認知再構成を行うため、非常に個人的な問題に関しては投稿をためらってしまう場合がある [14]。

チャットボットを使用したシステムでは、このような個人的な内容に関しても取り組むことができる。Fitzpatrick ら [12] は Woebot, Ly ら [21] は Shim というチャットボットを開発した。どちらも全自動のチャットボットであり、自己支援型の CBT プログラムを提供する。チャットボットを 2 週間使用した後、ユーザの不安、うつ、ストレスの度合いが低下したことが示された。Shinozaki ら [28] はソフトウェア開発者用のカウンセラーチャットボットを設計した。これは対話を通じてユーザの抱える問題を発見し、パターンマッチングによってカウンセリングの知識ベースから解決策を探索するものである。これらのチャットボットは、ユーザの話聞くことによって支援を行うことが可能であるが、ルールベースの設計のために柔軟性に欠けた応答が行われてしまうことがある [2]。

ユーザの話聞くという設計とは対照的に、Kim ら [19] は、ユーザ側がボットの支援を行うことの、ユーザ自身の心理状態への影響を検証した。著者らは Gloomy という名前のソーシャルボットを構築した。Gloomy はフェイスブック上に自身の悩みを投稿し、ユーザはその投稿に対して励ますような形で Gloomy を支援した。このやり取りにおいて、ユーザは Gloomy に対し客観的な視点から助言を行うことによって、自身が抱く悩みについても認知再構成を行うことができ、結果として心理状態が改善されることが示された。しかしながら、Gloomy の投稿内容は人手で作成されているため、このシステムの実用的な実用性は限定的である。さらに、研究の実験に比較対照群が設定されていないため、異なる種類のチャットボットがユーザの認知再構成および心理状態にどのような異なる影響を与えるかは未だ十分に明らかになっていない。

2.4 大規模言語モデル

近年性能が大きく向上した大規模言語モデルは、自然言語処理のさまざまなタスクで卓越した能力を発揮している [8, 24]。この進歩は、自然で柔軟な会話を行う全自動チャットボットを実現する上でも大きな光明であり [1]、異なる種類のチャットボットとの会話がユーザの認知再構成および心理状態にどのような異なる影響を与えるのかの理解が不足しているという研究ギャップを埋める機会を提供する。

そこで本研究では、大規模言語モデルを活用して、ユーザに支援を行うチャットボットおよびユーザから支援を受

けるチャットボットという異なる 2 種類のチャットボットを組み込んだシステムを構築し、それらがユーザの認知再構成および心理状態にどのような異なる影響を与えるのかを比較検証した。

3. システム構成

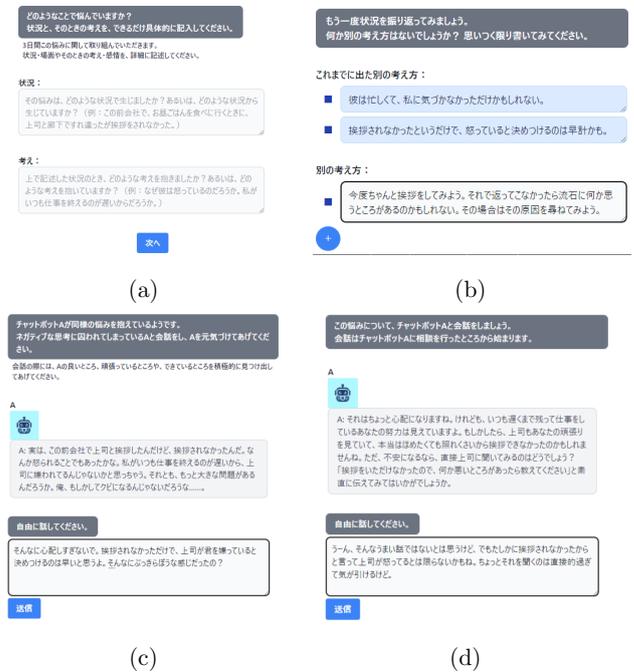


図 1: ウェブシステムのインターフェース。(a) 現在の悩みを記入するページ。(b) 別の考え方を記入するページ。(c) ピアチャットボットと会話をするページ。(d) コンサルティングチャットボットと会話をするページ。

3.1 全体の構成

本システムは Firebase*1 を用いて、ウェブベースで構築した。チャットボットとの会話には OpenAI の GPT-4 [24] を利用し、OpenAI API*2 を通じて GPT-4 からのレスポンスを取得した。

3.2 チャットボット設計

我々は 2 種類のチャットボットを設計した。一つはユーザと同様の悩みを語るピアチャットボットであり、もう一つはユーザの悩みに対して直接助言を行うコンサルティングチャットボットである。

ピアチャットボットは、心が弱く傷つきやすい (vulnerable) 考え方をするように設計した。具体的には、良い側面を過小評価し、悪い側面を過大評価したり、過度な一般化を行ったり、極端な考え方をするように設計した。自身と同様の悩みを語るチャットボットに対して第三者と

*1 <https://firebase.google.com/>

*2 <https://platform.openai.com/docs/overview>

して助言を行うことによって、ユーザが自身の状況を客観的に捉えなおし、より建設的な考え方を思いつけるような間接的支援が実現されることを狙った。

一方で、コンサルティングチャットボットは、心が強くくじけない (resilient) 考え方をするように設計した。具体的には、物事の良い側面を探し出し、改善・成長できる部分を見出すような思考をするように設計した。自身の悩みについてチャットボットに相談を行い、助言を受けることによって、別の視点からのより建設的な考え方を学べるような直接的支援が実現されることを狙った。

3.3 システムのインタフェース

図 1 に本ウェブシステムのインタフェースを示す。まず、ユーザは自身が現在抱えている悩みについて、その悩みが生じた状況とそのときの考えを入力する (図 1a)。次に、ユーザは先ほど記入した悩みに関して、チャットボットと会話を行う。ユーザは、自身が入力した悩みと同様の悩みを語るチャットボットの悩み相談に乗る (図 1c)、あるいはユーザ自身が入力した悩みに関してチャットボットに相談する (図 1d)。この会話はチャットボットがはじめ、チャットボットとユーザは交互に発言をする。ユーザが3回、チャットボットが4回発言したところで会話は終了する。その後、ユーザは再度悩みの生じた状況を振り返り、別の考え方を可能な限り入力する (図 1b)。

4. ユーザ実験

4.1 手順

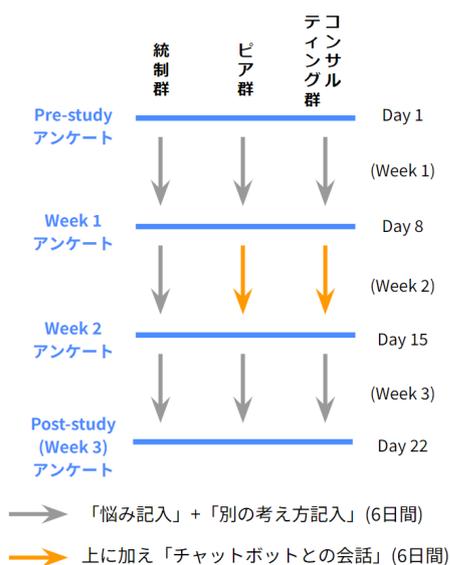


図 2: ユーザ実験の流れ。

図 2 にユーザ実験の流れを示す。1 日目には、オンボーディングを行った後、Pre-study アンケートを実施した。2 日目から 7 日目 (Week 1) には、全参加者は「悩みを記

入」しその直後に「別の考え方を記入」する、という作業を毎日行った。参加者の作業負担を考慮し、悩みの記入は 3 日ごとに行う設計とした。すなわち、参加者は 2, 3, 4 日目で一つと同じ悩みについて取り組み、5, 6, 7 日目で次の悩みについて取り組む、というような設計とした。8 日目には、Week 1 アンケートを実施した。9 日目から 14 日目 (Week 2) には、ピア群・コンサルティング群の参加者は「悩みを記入」したあとにチャットボットとの会話を行い、その後「別の考え方を記入」した。悩みの記入は Week 1 と同様に 3 日ごととした。統制群の参加者は Week 1 と同様の作業を行った。15 日目には、Week 2 アンケートを実施した。16 日目から 21 日目 (Week 3) には、全参加者が Week 1 と同様の作業を行った。22 日目には、Post-study アンケートを実施し、オフボーディングののち実験が終了した。

実験終了後、各参加者には 3200 円の謝礼を支払った。加えて、参加者 1 人あたり 903 円を手数料としてクラウドワークスに支払った。

4.2 評価尺度

| Pre-study | Week 1 | Week 2 | Post-study (Week 3) |
|-----------|---------------------|---------------------|---------------------|
| WHO-5 | WHO-5 | WHO-5 | WHO-5 |
| SDS | Qs. about Reframing | Qs. about Reframing | SDS |
| STAI | | Qs. about Chatbot | STAI |
| TIPI | | | Qs. about Reframing |
| | | | Qs. about System |

図 3: 各アンケートにおける質問項目の一覧。

4.2.1 心理状態の測定

参加者の心理状態を測定する尺度として、ウェルビーイング尺度である WHO-5 の日本語版 [4]、うつ症状尺度である SDS の日本語版 [31]、不安尺度である STAI の日本語版 [30] を採用した。STAI に関しては、一般的で長期にわたる不安レベルを評価するために、特性不安部分 (X-2) のみを使用した。

これらのアンケートの使用は、先行研究 [19] に倣って決定した。先行研究ではうつ症状尺度として BDI-2 [6] が使用されていたが、著作権保持者である Pearson は BDI-2 をオンライン形式で使用するオプションを提供していなかったため、使用しなかった。SDS と STAI のオンライン形式での使用は、日本の著作権保持者である三京房*3によって許可された。

4.2.2 認知再構成の実践に関する質問

心理状態の測定尺度に加え、毎週の終わりにその週の認知再構成の実践結果についての質問を行った。例として、以下に Week 1 アンケートでの質問項目を提示する。

(1) 「現在の悩みを書く」という作業の難易度は、どれほ

*3 <https://www.sankyobo.co.jp/index.html>

- どだと感じましたか？(SD法による7段階評価. 1側に「容易だった」、7側に「難しかった」が対応する)
- (2) 「別の考え方を書く」という作業の難易度は、どれほどだと感じましたか？(上記と同じ7段階評価)
- (3) 2~4日目で取り組んだ悩みについて、自分の中でこの悩みは解決しましたか？解決した場合は、何日目で解決しましたか？(選択肢: 「2日目で解決した」「3日目で解決した」「4日目で解決した」「解決しなかった」)
- (4) (上記と同じ質問. ただし日にちは5日目から7日目に対応する)

4.2.3 チャットボットに関する質問

Week 2 アンケートでは、ピア群・コンサルティング群の参加者に対して、チャットボットに関する追加の質問を行った。内容としては、チャットボットと参加者の会話が想定通りのものとなっていたか確認するシステムの挙動確認の質問や、チャットボットに対する印象を調べるための記述式質問等である。

また、Post-study アンケートでは、Week 1 と Week 3 を比較して、悩みへの向き合い方の変化や、全体の感想を尋ねる記述式質問を行った。

なお、以上の質問項目に加え、Pre-study アンケートではパーソナリティ尺度である TIPI の日本語版 [25] も測定したが、今回の分析には用いなかった。

4.3 参加者

クラウドソーシングサイト*4上で20代前半から60代前半までの102名の参加者を集めた(男性・女性とも51名ずつ)。年代および性別の偏りが小さくなるよう留意しつつ、各参加者を統制群・ピア群・コンサルティング群の3群にランダムに振り分けた。コンサルティング群の1名がWeek 2 アンケートの回答後にユーザ実験を離脱し、残りの101名の参加者は全実験内容を完了した。

4.4 安全対策

本ユーザ実験は、募集に先立ち、東京大学工学部環境安全管理室の倫理審査委員会からの承認を得た(承認コード: KE23-102)。システムおよびユーザ実験の設計を決定する際には、臨床心理士との協議のもと、参加者に過度の心理的負担をかけない設計となるよう注意した。さらに、研究の安全性を高めるため、参加者の募集時に極端なケースを除外した。具体的には、同意書に「精神疾患の既往歴がなく」「現在心理カウンセリングを受けていない」ことを確認する質問を設け、このような人が実験に参加しないようにした。この除外基準は、精神的に不安定な人の心理状態や、現在受けている心理カウンセリングの効果に対して影響を及ぼしてしまうリスクを避けるために設定した。これ

らの安全対策は、先行研究 [19] を参考に設定した。

5. 結果

以下に定量分析および定性分析を行った結果を示す。結果の分析には、統計分析用オープンソースソフトウェアの JASP*5を使用した。

5.1 定量分析結果

5.1.1 システムの挙動確認

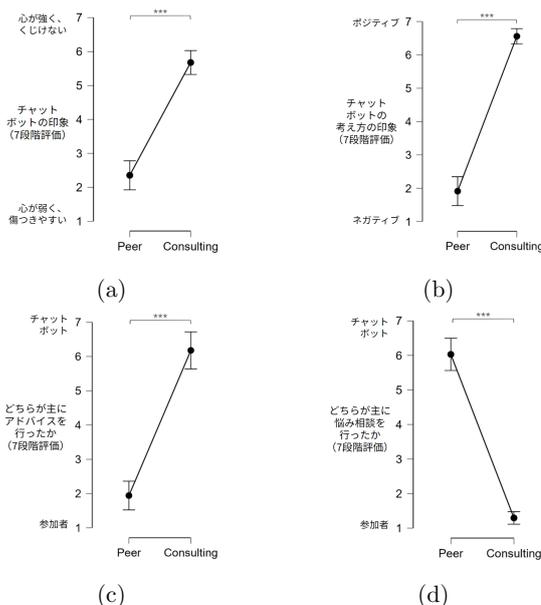


図 4: システムの挙動確認の結果。(a) チャットボットの印象, (b) チャットボットの考え方の印象, (c) どちらが主にアドバイスを行ったか, (d) どちらが主に悩み相談を行ったか, をそれぞれ7段階評価で参加者に回答してもらった。全回答結果について有意差が見られた (***) $p < .001$ 。

チャットボットが設計通りの挙動を行ったか確認するための質問の回答に対し t 検定を行った結果を図 4 に示す。Welch の t 検定の結果、想定通り参加者はコンサルティングチャットボットをピアチャットボットよりも心が強くくじけないと感じ (図 4a, $p < .001$)、コンサルティングチャットボットがピアチャットボットよりもポジティブな視点を持っていると感じた (図 4b, $p < .001$)。また、Welch の t 検定の結果、想定通り、コンサルティングチャットボットがピアチャットボットよりも悩みに対するアドバイスをを行ったと感じられ (図 4c, $p < .001$)、ピアチャットボットがコンサルティングチャットボットよりも悩みの相談を行ったと感じられた (図 4d, $p < .001$)。

5.1.2 心理状態の変化

各心理尺度の反復測定分散分析の結果を図 5 に示す。SDS, STAI, WHO-5 いずれの尺度においても、各週・各

*4 クラウドワークス <https://crowdworks.jp/>

*5 <https://jasp-stats.org/>

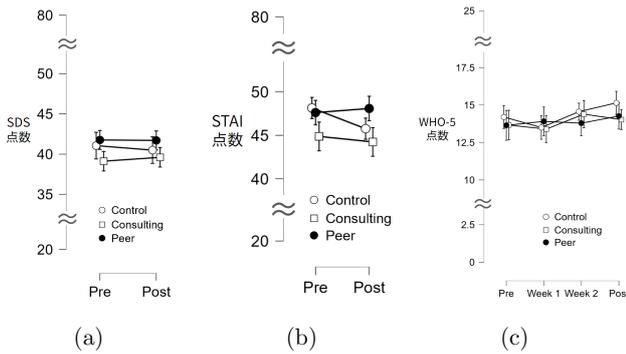


図 5: 心理状態の測定結果. 各参加者の (a) SDS の得点, (b) STAI の得点, (c) WHO-5 の得点, の変化を測定した.

群間に有意差は見られなかった.

5.1.3 「現在の悩みを書く」主観的難易度

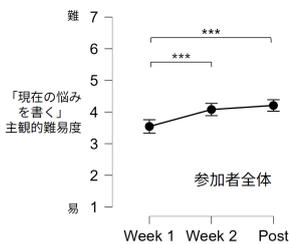


図 6: 「現在の悩みを書く」主観的難易度の 7 段階評価. 参加者全体では, Week 1 から Week 2 および Week 1 から Week 3 にかけて有意な増加が見られた (***) ($p < .001$).

「現在の悩みを書く」という作業の主観的難易度を反復測定分散分析した結果を図 6 に示す. 群別に見たときでは, 各週・各群間に有意差は見られなかったが, 参加者全体では, Week 1 から Week 2 および Week 1 から Week 3 にかけて主観的難易度が有意に増加した ($p < .001$).

5.1.4 「別の考え方を書く」主観的難易度

「別の考え方を書く」という作業の主観的難易度を反復測定分散分析した結果を図 7 に示す. 参加者全体では, 各週間に有意差は見られなかったが, 群別に見たときでは, ピア群において, Week 1 から Week 2 にかけて主観的難易度が有意に増加した ($p < .05$). すなわち, ピアチャットボットの介入によって, 別の考え方を書くという作業が有意に難しく感じられたことが確認された.

5.1.5 別の考え方の個数

各参加者が各週で記入した別の考え方の個数を反復測定分散分析した結果を図 8 に示す. 群別に見たときでは, 各週・各群間に有意差は見られなかったが, 参加者全体では, Week 1 から Week 2 にかけて別の考え方の個数が有意に減少した ($p < .01$).

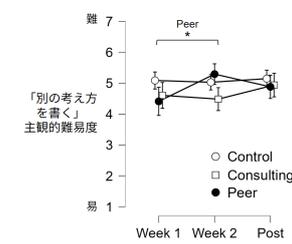


図 7: 「別の考え方を書く」主観的難易度の 7 段階評価. 群別に見たときでは, ピア群において, Week 1 から Week 2 にかけて有意な増加が見られた ($*p < .05$).

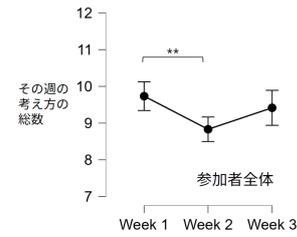


図 8: 各週で記入された別の考え方の個数. 参加者全体では, Week 1 から Week 2 にかけて有意な減少が見られた (** $p < .01$).

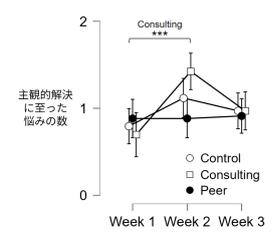


図 9: 各週で主観的解決に至った悩みの個数. 群別に見ると, コンサルティング群で, Week 1 から Week 2 にかけて有意な増加が見られた (***) ($p < .001$).

5.1.6 主観的解決に至った悩みの個数

各参加者が各週で記入した 2 つの悩みのうち, 解決に至ったと回答した個数を反復測定分散分析した結果を図 9 に示す. 群別に見ると, コンサルティング群で, Week 1 から Week 2 にかけて, 解決に至った悩みの個数が有意に増加した ($p < .001$). また, 参加者全体でも, Week 1 から Week 2 にかけて, 解決に至った悩みの個数が有意に増加した ($p < .001$). すなわち, コンサルティングチャットボットによる介入によって, 参加者は自身の悩みに対する解決感を強く感じられるようになったことが確認された.

5.2 定性分析結果

5.2.1 チャットボットについて

Week 2 アンケートでのチャットボットに関する記述式質問の回答結果を以下に述べる.

まず, ピアチャットボットに対して, ピア群に属する参加者 34 人中 12 人が, チャットボットを通して状況の客観視がしやすくなったと回答した.

対人関係の悩みでは, 自分と悩みの対象に, さらにチャットボットが加わることで, 自分が第三者として悩みを客観視しやすくなり解決の糸口を探しやすくなるように感じた. [P10]

あまりにもネガティブなので, 自分に置き換え, 考えすぎるのも良くないなと気づきにつながり役に立った. 自分の気持ちを前向きにするのに有効だと感じた. [P76]

一方で, 34 人中 14 人が, チャットボットのネガティブな応答を批判的に評価した.

一方, 悩みの理由が自分起因である場合, チャットボットの発言に共感してしまい, 悩みに対するネガティブな感情が補強されて, 落ち込みをよりひどくしたように感じる. [P10]

悪かったことは, 一時, こちらが何を提案してもネガ

ティブな考えが止まらなかった点です。もっと積極的に結果がどうであれ、あれこれ悩んでいる暇があったら自分でやって試してみたらよいのにと思いました。 [P34]

また、コンサルティングチャットボットに対して、コンサルティング群に属する34人中11人が、自分とは違う視点から解決策を提案してくれたことを肯定的に評価した。また、34人中14人が、とにかく共感して慰めてくれるような姿勢を肯定的に評価した。

相談者に対して、とてもポジティブに共感してくれて、褒め称えてくれるところは、ちょっと仰々しいな、わざとらしいなど思うこともありました。悪い気はしませんでした。普段の生活で、そんなに褒めてくれる人はいませんから。 [P98]

一方、34人中3人が、チャットボットからのアドバイスに具体性が欠けることを指摘した。また、2人が実際の心理としてできそうにないことをアドバイスされたことを指摘した。中には、チャットボットの共感に対し抵抗感を抱いた参加者もいた。

何事にも、まず共感というスタイルを貫いていました。しかし、その共感に問題があるようにおもえた。なぜ、AIが「わかります」なんて言うのか？お前に何がわかるんかい！とか感じた [P35]

6. 考察

6.1 ユーザの認知再構成に対するチャットボットの影響

6.1.1 悩みの記述の難易度

悩みを記述することに関してはチャットボットの影響は特に見られなかったが、参加者全体として、Week 1 よりも Week 2, Week 3 において自身の悩みを記述することがより難しく感じられることが確認された。これは、日常的に認識している具体的な悩みがそれほど多くないためだと考えられる。つまり、常に意識していた2, 3個の悩みに取り組んだ後、参加者は自身に関する新たな問題を見つけなければならなかった。その際、参加者は自身を深く省みて、抽象的な不満を言語化する必要があった。この具体化作業が、先行研究 [18] に示されているように、心理的な負担をもたらした可能性があり、結果として悩みを記述することがより難しく感じられたのだと考えることができる。

6.1.2 別の考え方の記述の難易度

先行研究 [19] からの予想に反し、ピアチャットボットによって参加者が別の考え方をより記述しやすくなるというような効果は測定されなかった。むしろ、ピア群ではチャットボットの介入によって別の考え方を記述することが有意に難しく感じられたことが確認された。この原因として、認知再構成法に関する参加者のマインドセットが十

分に整えられていなかった可能性が挙げられる。我々は参加者に認知再構成法や他者を支援する方法についての詳細な情報・説明を提供しなかったが、先行研究 [19] では、参加者がボットとの対話を行う前に約40分間の入門セッションが実施されていた。そのセッションでは、他者を支援する際のサポートガイドが提示された。具体的には、「不合理な思考の特定」「その思考の再構成」「社会的（感情的）支援の実施」の3つの手順についての説明がされた。これらの教育的知識の獲得、つまり、認知再構成法の実践手順や、他者に対する支援の行い方、他者を助けることによってもたらされる利益を理解することが、チャットボットを支援する際にユーザがその効果を最大限引き出す上で重要な要因となるのだと考えることができる。したがって、このようなチャットボットシステムを構築する際には、チャットボットとの会話前にユーザが認知再構成法および他者を支援することなどの前提知識を学ぶことができるようにし、十分な心構えができるよう設計すべきだと考えられる。

6.1.3 悩みの解決感

コンサルティング群においてのみ、チャットボットが介入した週において、参加者が解決したと感じた悩みの個数が有意に増加した。これは、コンサルティングチャットボットと会話をしたことによって、参加者が自身の悩みに対してより解決感を感じることができたということを示唆している。コンサルティングチャットボットからの助言が時に具体性や実現可能性に欠けると述べた参加者もいたものの、少なくない参加者がチャットボットのポジティブな発言によって助けられたと述べている。すなわち、チャットボットから継続的に臆面なく届けられる肯定的な言葉が、参加者を快く、安心し、励まされ、問題に立ち向かおうとする気分させ、それが悩みに対する解決感につながったのだと考えることができる。

6.2 ユーザの心理状態に対するチャットボットの影響

先行研究 [19] とは異なり、いずれの群においても、参加者の心理状態に有意な変化は見られなかった。その原因としては、先と同様に認知再構成法に関する参加者のマインドセットが十分に整えられていなかった可能性のほかに、一つの悩みに取り組む時間が短く、心理状態の改善が得られるには向き合い方が不十分であった可能性が挙げられる。特に、Week 1 で記入された悩みに関しては、悩みを記入するのが Week 2, 3 と比して容易であったという結果を踏まえると、以前から心に抱いていた重要な問題であった可能性がある。しかしながら、この悩みに関して取り組む期間は3日間のみであって、その短期間のうちに十分に解決されることは少なかったのではないかと考えられる。そのような重要な問題が未解決のまま心の中でくすぶり続けてしまったことで、心理状態の改善が見られなかった可能性が考えられる。したがって、今回のようなチャットボッ

トシステムの構築の際には、ユーザが十分だと感じられるまで、自分の取り組みたい悩みに取り組み続けられる設計が大切で、一つの問題に対して時間をかけて深く考えられるような設計とすることが重要だと考えられる。

7. 本研究の制約と今後の展望

本研究には複数の制約が存在する。まず、参加者の属性に偏りがある可能性がある。我々はクラウドソーシングサイトで参加者を募集した。クラウドソーシングで働く人々は、特に職業に関して属性の偏りがある可能性がある。様々な職業の参加者を含めることによって、より一般的な知見を得ることができると考えられる。

また、大規模言語モデルに与えたプロンプトが、必要十分な記述となっていたかの検証が十分になされていないという制約が存在する。システムの挙動確認は想定通りの結果であったものの、プロンプトの各部分がチャットボットの応答の品質にどう影響するかの検証を行うことで、認知再構成を支援するチャットボットシステムについてのより深い知見を獲得することができると考えられる。

さらに、ユーザ実験の設計に関しても制約が存在する。第一に、実験は3週間実施したが、参加者がチャットボット介入を体験したのは1週間のみであった。これは、参加者の心理状態に十分な影響を与えるには短すぎた可能性がある。また、参加者は2種類のチャットボットのうち1つとのみ会話ができただため、各参加者が2種類のチャットボットをどのように比較するかについての知見を得ることができなかった。したがって、今後の調査では、参加者がより長期間にわたって両種のチャットボットを体験できるユーザ実験を実施することで、各チャットボットに対するより深い洞察を行うことができると考えられる。

8. 結論

本研究では、認知再構成法における、異なる種類のチャットボットとのやり取りの効果を調査するため、大規模言語モデルを使用した2種類のチャットボットを組み込んだウェブシステムを開発した。ユーザ実験の結果、先行研究のような心理状態の改善は確認できず、ユーザに悩み相談をするチャットボットに関しても、予想に反し、ピアチャットボットの介入によって悩みに対する別の考え方を書くことが難しく感じられたことが確認された。これらの結果および先行研究の知見を踏まえ、認知再構成法を支援するチャットボットシステムの重要な構成要因について議論した。具体的には、システムの使用にあたってユーザの認知再構成に関するマインドセットを十分に整えること、一つの問題に対してもっと時間をかけて深く考えられるような設計とすることが重要だと考察した。今後の研究では、これらを意識的に設計したチャットボットシステムを開発し、ユーザの認知再構成にどのような影響をもたらす

かを調査すべきである。

謝辞 調査にご参加頂いた皆様と、本研究内容について貴重なご指摘を下さった矢谷研究室の皆様へ深く感謝いたします。

参考文献

- [1] Mahyar Abbasian, Iman Azimi, Amir M. Rahmani, and Ramesh Jain. Conversational health agents: A personalized llm-powered agent framework, 2024.
- [2] Alaa A. Abd-alrazaq, Mohannad Alajlani, Ali Abdallah Alalwan, Bridgette M. Bewick, Peter Gardner, and Mowafa Househ. An overview of the features of chatbots in mental health: A scoping review. *International Journal of Medical Informatics*, Vol. 132, p. 103978, 2019.
- [3] Stav Atir, Kristina A. Wald, and Nicholas Epley. Talking with strangers is surprisingly informative. *Proceedings of the National Academy of Sciences*, Vol. 119, No. 34, p. e2206992119, 2022.
- [4] Shuichi Awata, Per Bech, Sumiko Yoshida, Masashi Hirai, Susumu Suzuki, Motoyasu Yamashita, Arihisa Ohara, Yoshinori Hinokio, Hiroo Matsuoka, and Yoshimoto Oka. Reliability and validity of the japanese version of the world health organization-five well-being index in the context of detecting depression in diabetic patients. *Psychiatry and Clinical Neurosciences*, Vol. 61, No. 1, pp. 112–119, 2007.
- [5] Aaron T Beck. Cognitive therapy of depression, 1979.
- [6] Aaron T Beck, Robert A Steer, and Gregory Brown. Beck depression inventory-ii. *Psychological assessment*, 1996.
- [7] Charles S Carver and Michael F Scheier. Attention and self-regulation: A control-theory approach to human behavior, 2012.
- [8] Yupeng Chang, Xu Wang, Jindong Wang, Yuan Wu, Linyi Yang, Kaijie Zhu, Hao Chen, Xiaoyuan Yi, Cunxiang Wang, Yidong Wang, Wei Ye, Yue Zhang, Yi Chang, Philip S. Yu, Qiang Yang, and Xing Xie. A survey on evaluation of large language models, 2023.
- [9] David A Clark. Cognitive restructuring. *The Wiley handbook of cognitive behavioral therapy*, pp. 1–22, 2013.
- [10] Leslie F. Clark. Stress and the cognitive-conversational benefits of social interaction. *Journal of Social and Clinical Psychology*, Vol. 12, No. 1, pp. 25–55, 1993.
- [11] Bruce P. Doré, Robert R. Morris, Daisy A. Burr, Rosalind W. Picard, and Kevin N. Ochsner. Helping others regulate emotion predicts increased regulation of one's own emotions and decreased symptoms of depression. *Personality and Social Psychology Bulletin*, Vol. 43, No. 5, pp. 729–739, 2017. PMID: 28903637.
- [12] Kathleen Kara Fitzpatrick, Alison Darcy, and Molly Vierhile. Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (woebot): A randomized controlled trial. *JMIR Ment Health*, Vol. 4, No. 2, p. e19, Jun 2017.
- [13] Brandon A Gaudio. Cognitive-behavioural therapies: achievements and challenges. *BMJ Ment Health*, Vol. 11, No. 1, pp. 5–7, 2008.
- [14] Erving Goffman. Stigma: Notes on the management of spoiled identity, 2009.
- [15] Igor Grossmann and Ethan Kross. Exploring solomon's paradox: Self-distancing eliminates the self-other

- asymmetry in wise reasoning about close relationships in younger and older adults. *Psychological science*, Vol. 25, No. 8, pp. 1571–1580, 2014.
- [16] Stefan G Hofmann, Anu Asnaani, Imke JJ Vonk, Alice T Sawyer, and Angela Fang. The efficacy of cognitive behavioral therapy: A review of meta-analyses. *Cognitive therapy and research*, Vol. 36, pp. 427–440, 2012.
- [17] Tristen K Inagaki, Kate E Byrne Haltom, Shosuke Suzuki, Ivana Jevtic, Erica Hornstein, Julienne E Bower, and Naomi I Eisenberger. The neurobiology of giving versus receiving support: the role of stress-related and social reward-related neural activity. *Psychosomatic medicine*, Vol. 78, No. 4, p. 443, 2016.
- [18] Francis J. Keefe, Mark Lumley, Timothy Anderson, Thomas Lynch, and Kimi L. Carson. Pain and emotion: New research directions. *Journal of Clinical Psychology*, Vol. 57, No. 4, pp. 587–607, 2001.
- [19] Taewan Kim, Mintra Ruensuk, and Hwajung Hong. In helping a vulnerable bot, you help yourself: Designing a social bot as a care-receiver to promote mental health and reduce stigma. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, pp. 1–13, New York, NY, USA, 2020. Association for Computing Machinery.
- [20] Joël Billieux Lucien Rochat and Martial Van der Linden. Difficulties in disengaging attentional resources from self-generated thoughts moderate the link between dysphoria and maladaptive self-referential thinking. *Cognition and Emotion*, Vol. 26, No. 4, pp. 748–757, 2012. PMID: 21972982.
- [21] Kien Hoa Ly, Ann-Marie Ly, and Gerhard Andersson. A fully automated conversational agent for promoting mental well-being: A pilot rct using mixed methods. *Internet Interventions*, Vol. 10, pp. 39–46, 2017.
- [22] Kateri McRae, Bethany Ciesielski, and James J Gross. Unpacking cognitive reappraisal: goals, tactics, and outcomes. *Emotion*, Vol. 12, No. 2, p. 250, 2012.
- [23] Robert A. Neimeyer and Guillem Feixas. The role of homework and skill acquisition in the outcome of group cognitive therapy for depression. *Behavior Therapy*, Vol. 21, No. 3, pp. 281–292, 1990.
- [24] OpenAI. Gpt-4 technical report. 2023.
- [25] Atsushi Oshio, Abe Shingo, and Pino Cutrone. Development, reliability, and validity of the japanese version of ten item personality inventory (tipi-j). *Japanese Journal of Personality*, Vol. 21, No. 1, pp. 40–52, 2012.
- [26] Tom Pyszczynski and Jeff Greenberg. Self-regulatory perseverance and the depressive self-focusing style: a self-awareness theory of reactive depression. *Psychological bulletin*, Vol. 102, No. 1, p. 122, 1987.
- [27] Peter Salovey. Mood-induced self-focused attention. *Journal of personality and social psychology*, Vol. 62, No. 4, p. 699, 1992.
- [28] Tetsuo Shinozaki, Yukiko Yamamoto, and Setsuo Tsuruta. Context-based counselor agent for software development ecosystem. *Computing*, Vol. 97, pp. 3–28, 2015.
- [29] C. Estelle Smith, William Lane, Hannah Miller Hillberg, Daniel Kluver, Loren Terveen, and Svetlana Yarosh. Effective strategies for crowd-powered cognitive reappraisal systems: A field deployment of the flip*doubt web application for mental health. *Proc. ACM Hum.-Comput. Interact.*, Vol. 5, No. CSCW2, oct 2021.
- [30] 中里克治, 水口公信. 新しい不安尺度 stai 日本版の作成: 女性を対象とした成績. *心身医学*, Vol. 22, No. 2, pp. 107–112, 1982.
- [31] 福田一彦, 小林重雄. 自己評価式抑うつ性尺度の研究. *精神神経学雑誌*, Vol. 75, pp. 673–679, 1973.