

中山人間科学振興財団
2024 年度研究助成活動報告書（研究助成）

研究課題名：
Chat GPT を用いた AI 心理士によって well-being を高める社会を創造できるか？

伊角 彩
東京科学大学 政策科学分野

背景

Well-being は「幸福で肉体的、精神的、社会的すべてにおいて満たされた状態」と世界保健機構（WHO）によって定義され、医療分野だけではなくあらゆる分野で well-being の実現・向上が目指されている。世界的に見ると、国民皆保険制度のもと質の高い医療を受けることができ、紛争もなく社会保険制度が整っている日本は、肉体的・社会的に満たされた状態をつくりやすいと言える。しかし、精神疾患患者数は年々増加傾向で 2020 年には推計 600 万人を超え（厚生労働省, 2025）、国民の幸福度も G7 の中で最下位である（Helliwell et al., 2024）。つまり「幸福で」「精神的に満たされた状態」を実現できない社会構造が制度以外の部分で存在すると考えられる。

第一の原因として、日本文化に根強く残る、メンタルヘルスに関するネガティブなイメージ（スティグマ・偏見）の存在がある。それゆえ、具体的には次のような問題が生じていると考えられる：①メンタルヘルスの不調は本人の弱さや至らなさが原因だと考える傾向にある、②不調が顕在化するまで自分は大丈夫と考えメンタルヘルスの問題を「自分ごと」として捉えていない人が多い、③自分のメンタルヘルスに向き合う習慣がなくセルフケアができている人が少ない、④メンタルヘルスケアを受けることの心理的ハードルが高い（例：第三者に相談することは恥ずかしい、どう相談していいのかわからない、相談していることを誰かに知られたくない等）。

第二に、日本では「メンタルヘルスケア＝精神科・心療内科」のイメージが強く、メンタルヘルス不調が起きて初めてケアを受ける仕組みになっている。公認心理師・臨床心理士が提供するカウンセリングを受けることができる民間施設も少なく、利用しようとしても高額であったり、時間や場所が限られていたりすることも多い。これは、メンタルヘルスケアを提供する民間施設が多くメンタルヘルスケアがもっと身近な存在である欧米とは大きく異なる点であり、不調が顕在化する前にケアを受けることを妨げている。

本研究では、上記のようなメンタルヘルスケアを受けるハードルを下げるために、Chat GPT を用いた AI 心理士を開発する。Chat GPT を用いた AI によるカウンセリングに関する研究は増えつつあり、メンタルヘルスのリテラシーを向上する（心理教育）ツールとして有用であることが示唆されたり（Alanezi, 2024）、認知行動療法に基づく対話のルールやシナリオ、応答例があらかじめ設定された Chat GPT を活用することで、利用者の気分の改善、認知の変化などが認められたという報告がある（Izumi et al., 2024）。しかし、情報の正確性と信頼性、利用者が危機的状況にある場合の対応、人間の複雑な感情の理解、人間関係の構築などに課題が残る。加えて、精神疾患患者を対象とした研究が中心となっており、メンタルヘルス不調が潜在化している層を含めた健常者を対象にした研究はまだほとんどない。

そこで、本研究では AI 心理士により予防的にメンタルヘルスケアを受ける仕組みを構築し、多くの国民が「幸福で」「精神的に満たされた状態」である社会を創造できるかを探求する。具体的に本研究では、以下の 3 つの点について検証することを目的とする。

1. AI 心理士はメンタルヘルスケアを受ける心理的ハードルを下げるのか。
2. 人間ではない AI 心理士に相談することは、メンタルヘルス不調の予防に対して効果（気持ち
が楽になる、解決策の参考になった等）があるのか。
3. AI 心理士から人間による対応に切り替えることについて、倫理的・法律的・心理学的視点か
ら適切なタイミングがどこなのか。

方法

まず、Slack 上で、ChatGPT-4o を用いた AI 心理士 α 版を開発するにあたり、プロンプトの作成・最適化を行なった^{注 1}。AI 心理士においては、AI を活用することで相談することの心理的ハードルを下げるだけでなく、AI が心理士のように、対等な立場で傾聴を重視し温かく寄り添う姿勢を保つことで AI 心理士との会話によってウェルビーイングが向上することを目指した。そのため、心理士の専門的知見を得ながらプロンプトの作成を進めた。また、パーソナライズされた会話ができるよう、会話開始時に、援助希求への態度（人に相談することが好きかどうか）、口調（丁寧語/タメ口）や会話の長さの好みを尋ねるようにした。Slack 上で α 版が完成した後、10 名の一般利用者を対象にテスト運用を行い、利用した感想などフィードバックを得た。

次に、LINE で実装するためにプロンプトの最適化を行い、AI 心理士 β 版を作成した。LINE アカウント「B-self」を開設し、共同研究者や心理士とともに動作確認を行なった。心理士からは β 版の改善点について知見を伺い、プロンプトに反映し精度を高めた。この AI 心理士 β 版を最終版（例は下記写真の通り）として効果検証を行うため、次の通り研究を行なった^{注 2}。なお、共同研究者と協議の上、本研究では AI 心理士から人間による対応には切り替えず、AI との会話がメンタルヘルスケアにつながるかを検証することとした。希死念慮や自殺企図を示すメッセージが届いた場合のみ緊急連絡先として厚生労働省によるホットライン等の案内を伝えることにした。

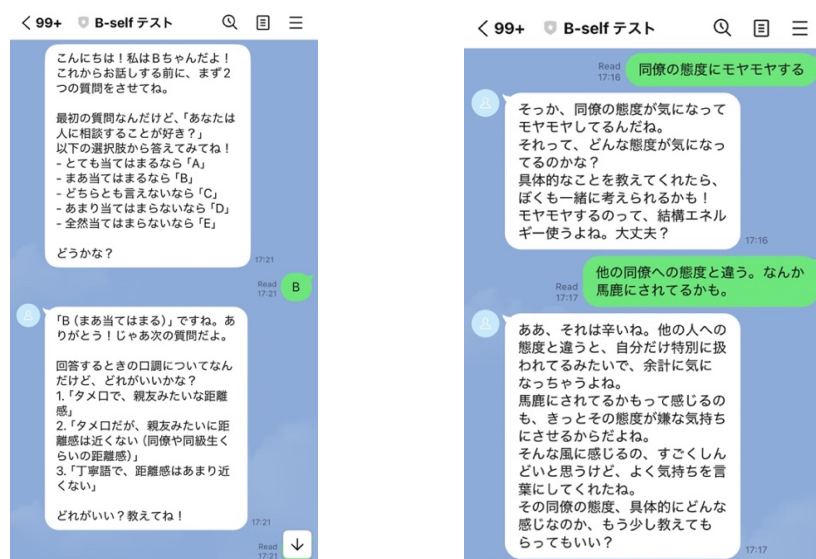


図1 LINE アカウント B-self 会話例

1. 対象者およびリクルート：就業している成人男女を対象として、スノーボールサンプリング（雪だるま式標本法）を用いてリクルートを行なった。メールや LINE 等を通じて研究説明および利用前アンケートへの URL を含む周知文を送り、研究参加に関心がある方には利用前アンケート（オンライン）に回答するように依頼した。利用前アンケートの最初に研究参加への同意欄を設けた。
2. パイロット実施方法：研究参加に同意し利用前アンケートに回答した研究参加者に対してメールを送り、B-self の利用方法を案内した。パイロット期間は平日・週末を含む 6 日間として、参加者の好きなタイミング・頻度で B-self を利用するよう勧めた。なお、より効果的に AI と会話ができるように以下の 3 つのポイントを伝えた。

- ・ B-self は、与えられた情報の元でアドバイスを提供することを優先する傾向にあります。まずは「話を聞いてほしい」「共感してほしい」「寄り添って欲しい」などといった要望がある方は、最初にその旨をチャット内でお伝えください。
例：●で困っている。まずは、しっかり話を聞いて寄り添ってから、必要なアドバイスが欲しい。
- ・ アドバイスが欲しい時には、状況のヒアリングや、問題・原因の深掘りをした上でアドバイスを提供するように指示してください。
例：アドバイスを提供するときは、私の状況の詳細をしっかり掘り下げた上でしてほしい。
- ・ あなたご自身が AI チャットに対して求める反応を先に、できるだけ明確にお伝えいただくと快適に会話をしていただきやすくなります。どんな反応を求めているか明確ではないときは、まずご自由に会話を始めてください。そして、あなたが求めている反応があった時には、その旨を指摘し、どういう反応をして欲しいか伝えてください。
例：そのアドバイスは私に合っていない。/複数の提案が欲しい。/アドバイスが薄っぺらいからもっと深掘りしてほしい。/共感するだけだと物足りない。どうすればいいか教えて。

3. 効果検証方法：利用前と利用後にオンラインアンケート（所要時間：各 10 分程度）を実施した。利用前アンケートでは、利用者の属性（性別、年齢、就業状況）、ウェルビーイング（主観的健康感、幸福度、自分への満足度、精神的健康度、抑うつ、不安、孤独感）、援助希求態度（困ったり悩んだりした時に相談する相手、実際に相談した相手）を尋ねた。なお、精神的健康度は WHO-5、抑うつは PHQ-2（Patient Health Questionnaire-2）、不安は GAD-2（Generalized Anxiety Disorder-2）、孤独感は日本語版 UCLA 孤独感尺度を用いて測定した。利用後アンケートでは、これらに加えて B-self 利用状況や利用に対する満足度・所感を調査した。

利用前・利用後アンケートにおいて、利用者および非利用者のウェルビーイングおよび

援助希求態度がどのように変化したかを比較し、B-self の利用によってウェルビーイングが高まるかを評価した。なお、サンプル数が少なく正規分布でなかったため Mann-Whitney U 検定を用いて、利用群と非利用群で、各指標の利用前の点数、利用後の点数、利用前後の変化量に差があるかを検証した。

注¹ AI 心理士の開発は共同研究者（土井理美）が所属する株式会社 BASNCO-SO が主体となって行なった。研究代表者は、心理士からの専門的知見や生成 AI のソフトウェア開発・研究開発を専門に行う会社のコンサルティングを受けながら、ウェルビーイングを高める会話を実現させるためのプロンプトの開発・最適化およびその効果検証を行なった。

注² 個人情報配慮のため、データ解析者（研究代表者）は個人情報を除いたデータを扱った。そのためアンケートの送付およびアンケートデータの管理は共同研究者（土井理美）が行なった。

結果

1. 参加状況

45 名が研究参加に同意し、利用前アンケートに回答した。そのうち、実際に B-self を登録・利用した方は 23 名（参加者における利用率 **51%**）、4 名（9%）は登録のみにとどまった。B-self の登録・利用に関わらず研究参加者全員に回答を依頼した利用後アンケートには 23 名（51%）が回答し、そのうち 1 度でも B-self を利用した方は 18 名（40%、B-self 利用者のうち 78%）であった。

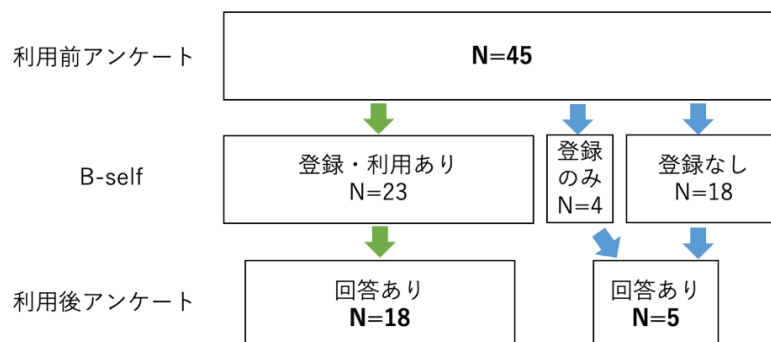


図 2 研究参加者フローチャート

2. 研究参加者の属性

研究参加者（N=45）は、女性が多く、年齢は 35～39 歳が最も多かった。就業している方を中心にリクルートしたため、フルタイムで働く方が中心であった。

研究参加者を、利用後アンケート未回答者（N=22）と回答者（N=23）、さらに回答者を B-self 利用群（N=18）と非利用群（N=5）に分けて属性を比較した。利用群と非利用群で属性に統計的な有意差は見られなかった（サンプル数が小さいため Fisher の正確確率検定を用いた）。

表 1 研究参加者の属性

		利用後アンケート								
		全体（N=45）		回答なし（N=22）		回答あり（N=23）				p*
						利用群（N=18）		非利用群（N=5）		
								人数	%	
性別		人数	%	人数	%	人数	%	人数	%	
	女性	31	68.9	15	68.2	13	72.2	3	60.0	0.621
	男性	14	31.1	7	31.8	5	27.8	2	40.0	
年齢	20歳未満	1	2.2	1	4.5	0	0	0	0	0.295
	25～29歳	3	6.7	2	9.1	1	5.6	0	0	
	30～34歳	2	4.4	1	4.6	1	5.6	0	0	
	35～39歳	14	31.1	4	18.2	9	50.0	1	20.0	
	40～44歳	4	8.9	2	9.1	1	5.6	1	20.0	
	45～49歳	8	17.8	4	18.2	3	16.7	1	20.0	
	50～54歳	3	6.7	1	4.6	1	5.6	1	20.0	
	55～59歳	5	11.1	3	13.6	2	11.1	0	0	
	60歳以上	4	8.9	3	13.6	0	0	1	20.0	
	答えたくない	1	2.2	1	4.6	0	0	0	0	
	就業形態	フルタイム	36	80.0	20	90.9	13	72.2	3	60.0
パートタイム		4	8.9	1	4.6	1	5.6	2	40.0	
専業主婦/主夫		2	4.4	0	0	2	11.1	0	0	
学生		3	6.7	1	4.6	2	11.1	0	0	

*Fisherの正確確率検定を用いて利用群と非利用群の分布に差があるか検定

3. B-self 利用状況・満足度等

B-self 利用者に対して、パイロット期間（6 日間）の利用状況や満足度等を利用後アンケートで尋ねたところ、以下のような結果となった。

(1) 利用状況

利用後アンケートに回答した利用者（18 名）のうち半数は 1 日のみの利用で、6 日間毎日利用したのは 1 名（5.6%）のみだった。また、1 日のやり取りの数は、8 割以上の方が 3 往復以内と短い会話であったことがわかった。

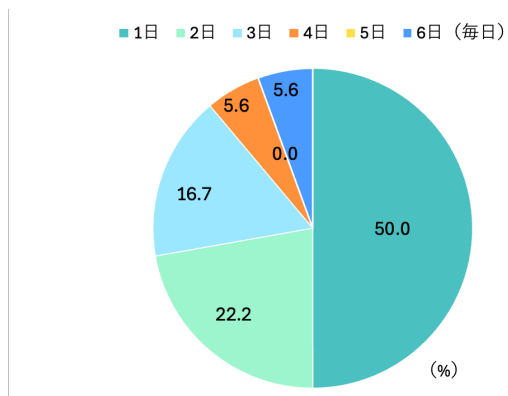


図 2 利用日数

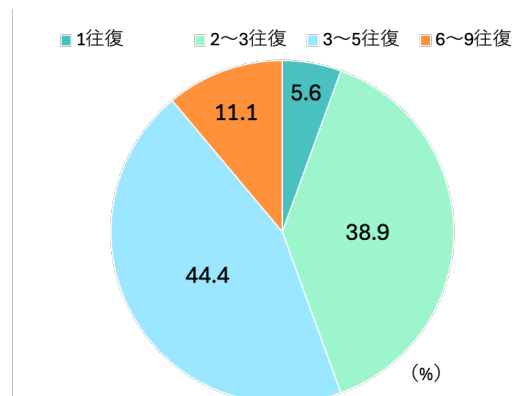


図 3 1日のやり取り数

利用するタイミングとしては、「自宅に1人で過ごしている時」が最も多く 44.4%、次に「通学・通勤中」(27.8%)が続いた。どのような時に利用することが多かったかを尋ねたところ、6割以上の方が「困りごとや悩みの解決策が知りたい時」に B-self を利用したと回答した。「何かモヤモヤすることがあった時」という回答も 55.6%と多かった。

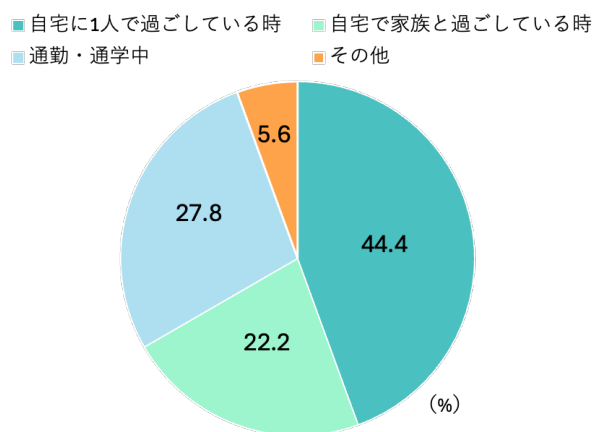


図4 利用のタイミング

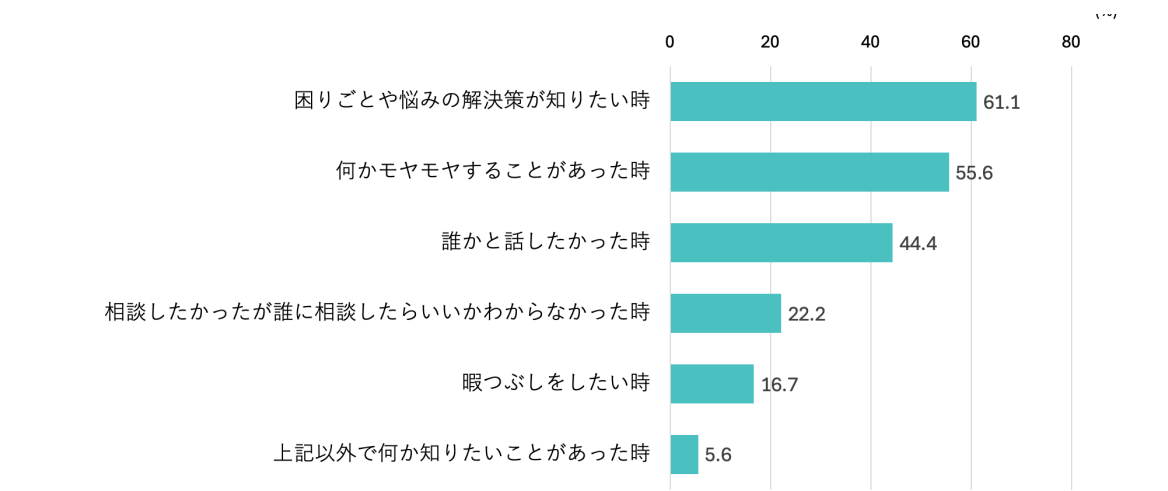


図5 どのような時に利用したか（複数回答可）

また、話した内容については、図6の通り、仕事に関することが最も多く (44.4%)、次に「職場の人間関係に関すること」「自分のメンタルヘルスに関すること」「雑談」(それぞれ 27.8%)が続いた。

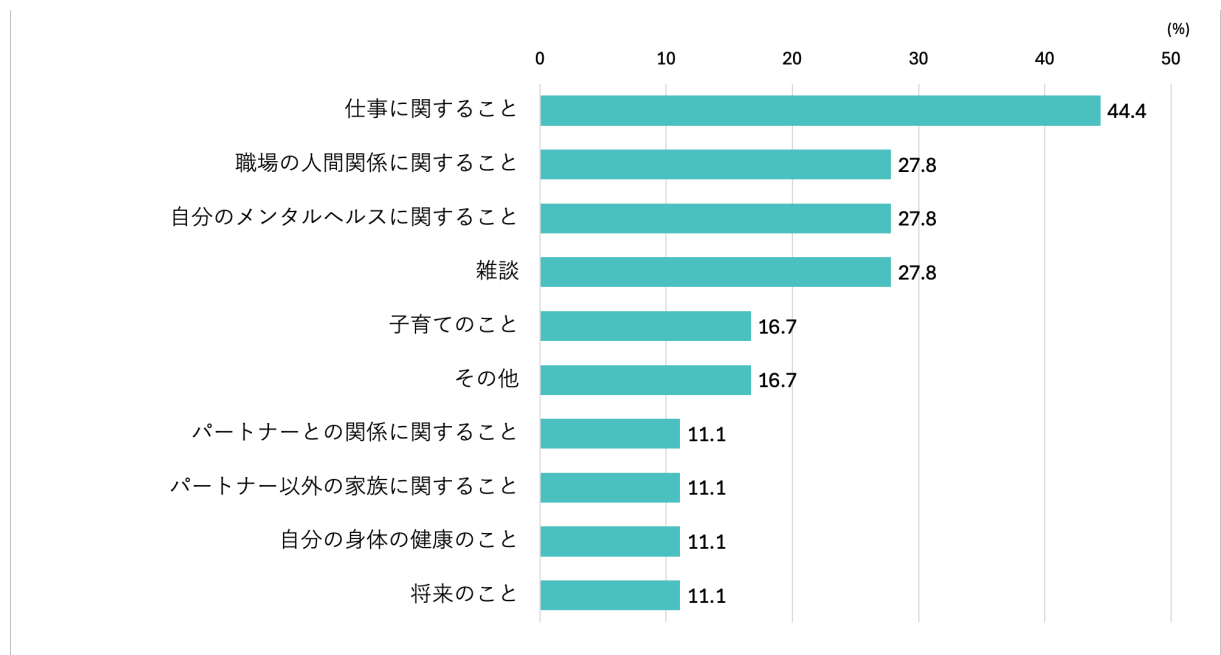


図6 話した内容（複数回答可）

(2) 利用満足度等

B-self 利用に対する満足度や所感を 10 つの指標を用いて測定したところ、以下のような結果となった。

① 利用満足度

利用満足度を 4 段階（とても満足している、まあ満足している、あまり満足していない、まったく満足していない）で尋ねたところ、27.8%が「とても満足している」、50.0%が「まあ満足している」という結果になった。

■ とても満足している ■ まあ満足している ■ あまり満足していない

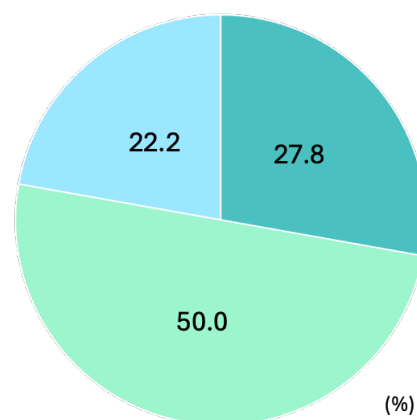


図7 利用満足度

満足している理由としては「不安な時や心落ち着かないとき、応援が欲しい時にコメントをもらえた」「相談相手になってもらえて、モヤモヤがはれた」「とても寄り添った返答で、よかった」「具体的なアドバイスをくれた」「悩んでいることをプラスの視点で解釈してくれて、励ましてくれるので、元気が出る時があった」などが挙げられ、寄り添った返答や具体的な解決策、ポジティブなフィードバックが満足度の高さに繋がったことがわかる。

一方、あまり満足していない理由としては「回答が薄っぺらく感じた」「返答が、機械的に感じた」「回答がワンパターン」など、表面的な回答が課題として残る。

なお、利用日数による満足度の差は見られなかった（結果省略）。

② 気軽に利用できたか

66.7%の方が「とてもそう思う」と回答し、気軽に利用できた方が多いことがわかった。

③ 話したかったことを十分に話すことができたか

「十分に話せた」または「まあ話せた」と回答した方は 83.3%に上った。

④ 気持ちが楽になったか

B-self を利用したことで気持ちが楽になった方は「とてもそう思う」「まあそう思う」を合わせると 72.2%であった。

⑤ ストレスケアができたか

B-self を利用したことでストレスケアができたかという問いに対して「とてもそう思う」と回答した方は 16.7%であり、「まあそう思う」と回答した方と合わせると 7 割近くに上った。

⑥ 安心感があったか

B-self と話すことに対して安心感があったかどうかについては、「とてもそう思う」「まあそう思う」を合わせると 88.9%となった。

⑦ 信頼できたか

B-self を信頼できたかどうかは「とてもそう思う」「まあそう思う」を合わせると上記と同じく 88.9%であったが、「とてもそう思う」の割合は安心感に比べて低かった。

⑧ 今後の対応方法や問題解決の参考になったか

この項目に対して「あまりそう思わない」と回答した方の割合は、他の項目と比べてもっとも多く、38.9%であった。利用状況を踏まえると、やり取りが短く利用日数も少ない方が多かったので、そういった状況では今後の対応方法や問題解決の参考になることは難しいことが示された。

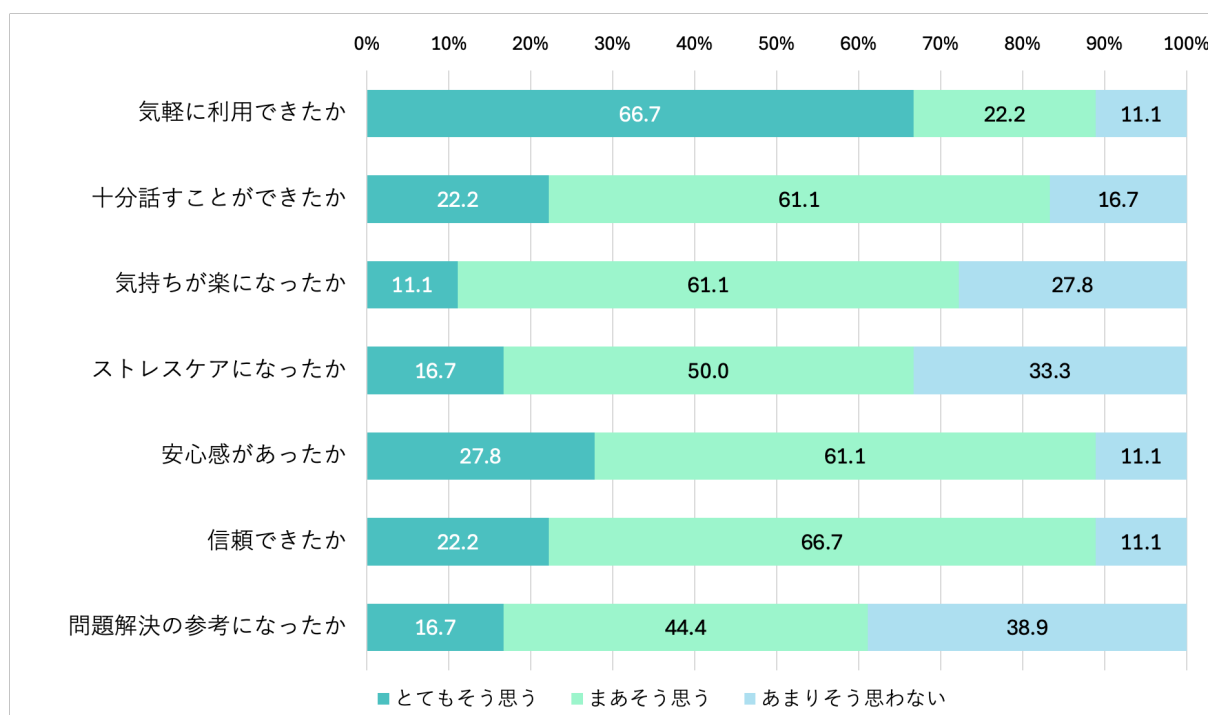


図8 利用に対する所感

⑨ 親しい友人や同僚にすすめようと思うか

B-self を親しい友人や同僚にすすめようと思うかを 0～10 点で尋ねたところ (Net Promoter Score: NPS®)、図9のような結果が得られた。平均は 6.83 点 (SD: 2.45 点) であった。10～9 と回答した方 (推奨者と定義される) は 27.8%、8～7 と回答した方 (中立者) は 27.8%、6～0 (批判者) 44.4%と、批判者が推奨者より多い結果となった。

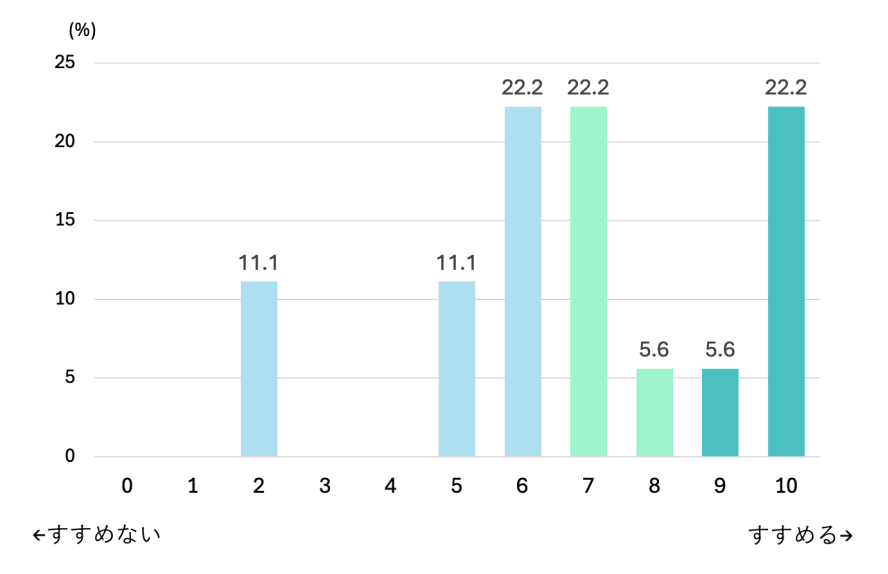


図9 親しい友人や同僚にすすめようと思う程度

⑩ 今後も B-self を利用したいか

利用者本人が今後も B-self を利用したいかについては、3 分の 1 の方が「とてもそう思う」と回答し、「まあそう思う」と回答した方を合わせると 83.3%となった。加えて、この方々に、月額いくらだと利用したいと思うか（Willing To Pay）を尋ねたところ、平均値は月額 220 円、中央値は 150 円、モード（最頻値）は 300 円（0 円を除いた場合）であった。なお、今後も利用したい方のうち 3 分の 1 は、有料の場合は使いたくない（0 円）と回答した。

■ とてもそう思う ■ まあそう思う ■ あまりそう思わない

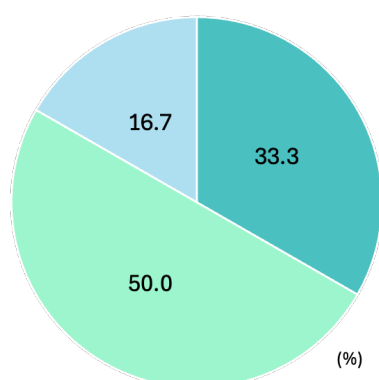


図 10 今後も利用したいか

表 2 Willing To Pay（月額）の分布

	人数	%
0円	5	33.3
100円	2	13.3
150円	1	6.7
250円	1	6.7
300円	4	26.7
500円	1	6.7
1000円	1	6.7
合計	15	100

今後 B-self をあまり利用したいと思わない理由には「chatGPT や自身による悩み解消策で間に合っているから」「対面で自分の信頼のできる人に相談したいタイプだから」「回答にもっと柔軟性がほしいから」といった回答があった。

4. 援助希求態度の変化

困ったり、悩んだりした時の相談する相手として回答した割合が利用前アンケートから利用後アンケートでどのように変化したかを利用群・非利用群で示すと図 11・図 12 のように、どちらの群においてもオンライン相談やチャット相談を利用する人は増えなかったが、父母や医療機関に相談する割合が増加したことがわかる。利用群においては、親戚、パートナー、職場の相談窓口相談する割合も増加していた。しかし、これらの利用前後での変化量は Fisher の正確確率検定の結果、両群とも統計的に有意な値ではなかった上、利用群と非利用群の差も見られなかった。

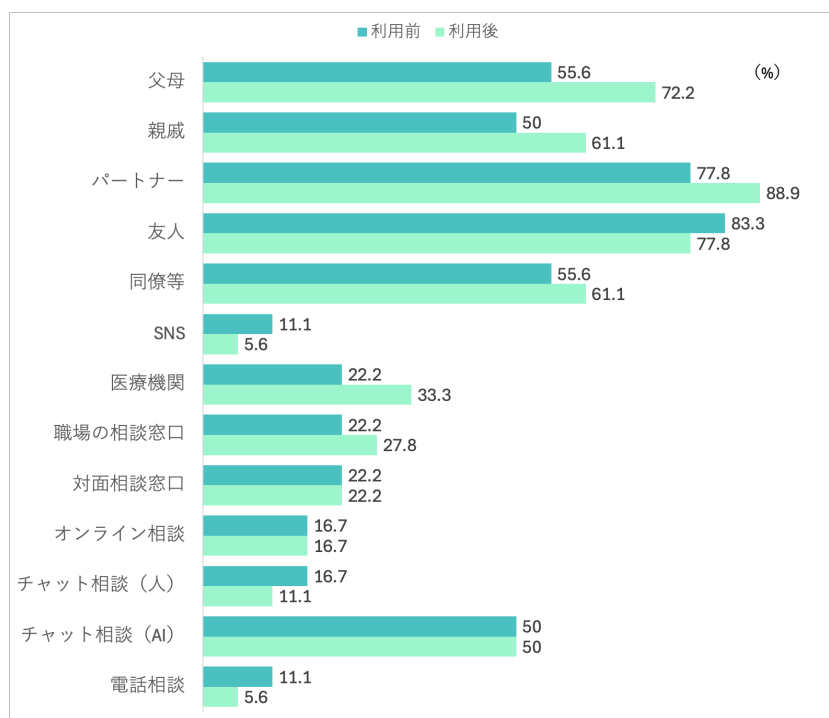


図 11 利用群 (N=18) における相談相手の変化

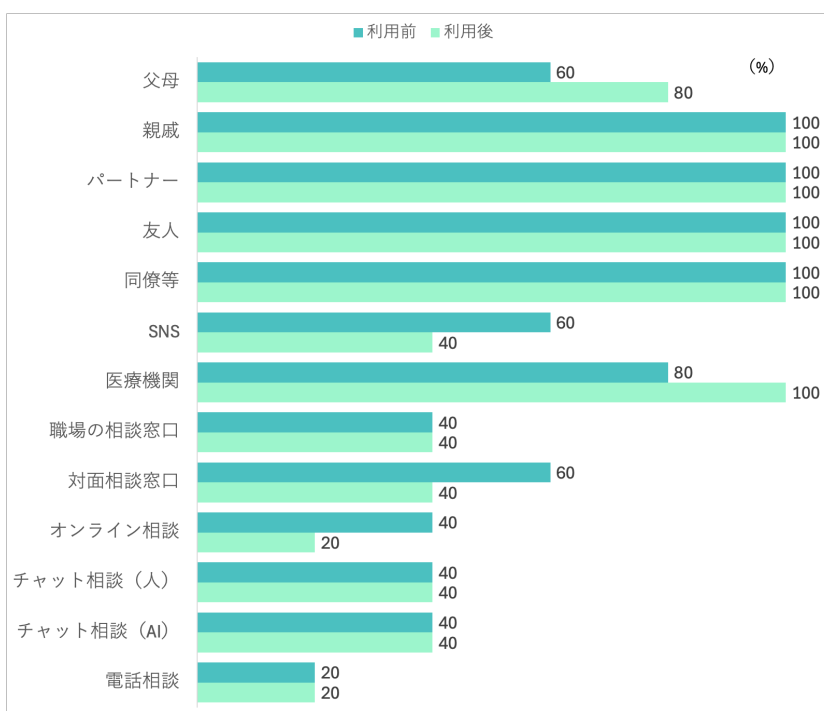


図 12 非利用群 (N=5) における相談相手の変化

5. ウェルビーイングの変化

Mann-Whitney U 検定を用いて、利用群と非利用群で、各指標の利用前の点数、利用後の点数、利用前後の変化量に差があるかを検証したところ次のような結果となった（平均値と標準

偏差 (SD) は参考として記載)。

(1) 主観的健康感

現在の健康状態を「良い」「まあ良い」「ふつう」「あまり良くない」「良くない」の5段階で回答を求めた (1~5 点、点数が高いほど健康状態が良い)。平均値をみると利用群は B-self 利用後に 0.11 点上がっていたが、Mann-Whitney U 検定の結果、利用群と非利用群で有意な差は認められなかった ($z = -0.81, p = 0.42$)。

(2) 幸福度

自分が幸せと思うかどうかを 0 点 (とても不幸せ) ~10 点 (とても幸せ) で尋ねたところ、利用前・利用後ともに非利用群の方が幸福度が高かったが有意差はなかった (利用前: $z = 1.12, p = 0.26$; 利用後: $z = 1.33, p = 0.18$)。また、変化量に関しても 2 群で有意な差は認められなかった ($z = 0.69, p = 0.49$)。

(3) 自分への満足度

今の自分にどのくらい満足しているかを 0 点 (全然満足していない) ~10 点 (大満足) で尋ねたところ、利用群・非利用群ともに利用後に満足度が下がる傾向が見られたが、利用前のスコア、利用後のスコア、変化量いずれにおいても 2 群で有意な差は認められなかった。

(4) 精神的健康度

精神的健康度を WHO-5 を用いて測定した (0~25 点、点数が高いほど精神的健康状態が良いことを示す) が、利用前・利用後どちらにおいても非利用群の方が精神的健康度が高かったが有意差は見られなかった (利用前: $z = 1.35, p = 0.18$; 利用後: $z = 1.46, p = 0.14$)。変化量に関しても 2 群で有意な差は認められなかった ($z = -0.23, p = 0.82$)。

(5) 抑うつ

PHQ-2 を用いて抑うつ傾向を評価した (0~8 点、点数が高いほど抑うつ傾向が強いことを示す) ところ、利用前は、利用群の方が非利用群より統計的に有意に点数が高く、抑うつ傾向が強かったことが示された ($z = -2.30, p = 0.02$)。利用後は 2 群の点数の差は小さくなったが有意差は認められなかった。平均値をみると、利用後アンケートで、非利用群は 0.20 点上がっていたのに対し利用群は 0.39 点下がっていたが、Mann-Whitney U 検定の結果、変化量に関しても 2 群間で有意な差は見られなかった ($z = 1.07, p = 0.28$)。

(6) 不安

不安傾向について GAD-2 を用いて評価した（0～8 点、点数が高いほど不安が強いことを示す）ところ、抑うつ同様、利用前は利用群の方が非利用群より統計的に有意に点数が高く、不安傾向が強かったことが示された（ $z = -2.55, p = 0.01$ ）。利用後は 2 群の点数の差は小さくなったものの有意差は認められなかった。平均値をみると利用群は B-self 利用後に 0.17 点下がっていたが、変化量も 2 群間で有意な差は認められなかった（ $z = 0.61, p = 0.54$ ）。

(7) 孤独感

日本語版 UCLA 孤独感尺度を用いて孤独感を測った（3～12 点、点数が高いほど孤独感が強いことを示す）結果、利用群・非利用群ともに利用後に孤独感が増す傾向が見られたが、利用前のスコア、利用後のスコア、変化量いずれにおいても 2 群で有意差は認められなかった。

上記をまとめると、利用前のウェルビーイングの状態があまり良くない、特に抑うつと不安が高い傾向にある方が B-self を利用したことが明らかになった。B-self 利用で抑うつが軽減する傾向は見られたものの、利用前後の変化量は利用群・非利用群ともに小さく、またサンプル数も少ないことから、B-self 非利用群と比べて利用群においてウェルビーイングが統計的に有意に向上するという結果は得られなかった。

表3 利用前後のウェルビーイングの変化

		利用前					利用後					変化量				
		平均値	SD	中央値	z値	p値*	平均値	SD	中央値	z値	p値*	平均値	SD	中央値	z値	p値*
主観的健康感 (1～5)	利用群 (N=18)	2.89	0.90	3	1.15	0.25	3.00	0.77	3	0.51	0.61	0.11	0.90	0	-0.81	0.42
	非利用群 (N=5)	3.40	0.55	3			3.20	1.10	4			-0.20	0.84	0		
幸福度 (0～10)	利用群 (N=18)	8.17	1.42	8	1.12	0.26	7.89	1.68	8	1.33	0.18	-0.28	0.83	0	0.69	0.49
	非利用群 (N=5)	9.00	1.41	10			9.00	1.41	10			0.00	0.00	0		
自分への満足度 (0～10)	利用群 (N=18)	7.06	1.63	7.5	0.19	0.85	6.94	1.59	7	0.15	0.88	-0.11	1.64	0	-0.09	0.93
	非利用群 (N=5)	7.40	2.51	7			7.20	2.28	7			-0.20	0.45	0		
精神的健康度 (WHO5, 0～25)	利用群 (N=18)	14.06	4.26	15	1.35	0.18	14.94	4.70	17	1.46	0.14	0.89	2.81	0.5	-0.23	0.82
	非利用群 (N=5)	18.20	6.69	22			18.80	6.38	21			0.60	1.67	1		
抑うつ (PHQ-2, 0～8)	利用群 (N=18)	1.56	1.42	1	-2.30	0.02	1.17	1.69	0	-0.58	0.56	-0.39	1.24	0	1.07	0.28
	非利用群 (N=5)	0.20	0.45	0			0.40	0.55	0			0.20	0.45	0		
不安 (GAD-2, 0～8)	利用群 (N=18)	1.72	1.45	1.5	-2.55	0.01	1.56	1.79	1	-1.75	0.08	-0.17	0.99	0	0.61	0.54
	非利用群 (N=5)	0.20	0.45	0			0.20	0.45	0			0	0	0		
孤独感 (UCLA, 3～12)	利用群 (N=18)	5.00	1.78	5	0.46	0.65	5.67	2.22	5.5	0.08	0.94	0.67	2.06	0.5	-0.76	0.45
	非利用群 (N=5)	5.40	1.95	6			5.60	1.52	6			0.20	1.64	0		

*Mann-Whitney U検定

考察

本研究では、Chat GPT を用いて、利用者に温かく寄り添うことができるように開発した AI 心理士がメンタルヘルスケアを受ける心理的ハードルを下げ、利用者のウェルビーイングを高めるかどうかを検証した。

メンタルヘルスケアを受ける心理的ハードルを下げるかについては、利用者のうち 9 割近くが「気軽に利用できた」と回答しており、他の相談手段より利用のハードルが低い可能性が示唆された。本研究や AI 心理士へのコメント（自由記述欄）でも気軽さを評価するものが複数あった。また、相談相手に関して尋ねた質問に関しては、AI が対応するチャット相談を相談相手として選んだ割合は利用前後で変わらなかったものの、パートナーや親戚、同じ職場の人、職場の相談窓口相談すると回答した方は利用群において増加傾向にあった。AI 心理士と会話をすることで、周囲に相談できる人がいるかを見直したり援助希求行動を促されたりした可能性はあるが、本研究においては変化量や利用群・非利用群の差に統計的有意性は認められず、その可能性を検証することはできなかった。

ウェルビーイングについては、利用前に抑うつや不安が高い傾向にある方が AI 心理士を利用しやすく、それが利用後に軽減する傾向は見られたが、非利用群と比べてウェルビーイングが統計的に有意に向上するという結果は本研究からは得られなかった。しかし、利用者の 7 割近くが AI 心理士とのチャットによって気持ちが楽になった、ストレスケアができたと回答しており、また、本研究や AI 心理士へのコメント（自由記述欄）にも「ちょっとしたモヤモヤや悩みを解消するにはとても良い」「寄り添いや、労い、適度な提案もあって、有効と思う」とあったことから、AI 心理士との会話がメンタルヘルス不調の予防に効果的だと感じる人は少なくないと考えられる。

今後は、対象者を増やしたり AI 心理士の利用期間を長くしたりした上で、AI 心理士が心理的ハードルを下げるか、ウェルビーイングを高めるかを再検討する必要があるが、本研究で「気軽に相談でき、温かく寄り添ってくれる AI 心理士」を実装できたことは大きな意義があると考えられる。ChatGPT など生成 AI を活用したメンタルヘルスケアは、人間といかにお互いの特性・利点を活かしながら連携・協働するかが大切である（古川・吉永 2024; Praxeotech Publishing, 2025）ことから、今後は、本研究で得られた AI 心理士の利点と限界を踏まえつつ、人間によるメンタルヘルスケアと効果的に組み合わせて、ウェルビーイングを高める方法を検討していく必要がある。

謝辞

本研究にご協力いただいた研究参加者の皆様に心から感謝申し上げます。

引用文献

厚生労働省（2025）『精神保健医療福祉の現状等について』（第4回精神保健医療福祉の今後の施策推進に関する検討会、参考資料4）

<https://www.mhlw.go.jp/content/11121000/001374464.pdf>

Helliwell, J. F., Layard, R., Sachs, J. D., De Neve, J.-E., Aknin, L. B., & Wang, S. (Eds.). (2024). *World Happiness Report 2024*. University of Oxford: Wellbeing Research Centre.

Alanezi, F. (2024). Assessing the effectiveness of ChatGPT in delivering mental health support: A qualitative study. *Journal of Multidisciplinary Healthcare*, 17, 461–471.

Izumi, K., Tanaka, H., Shidara, K., Adachi, H., Kanayama, D., Kudo, T., & Nakamura, S. (2024). Response generation for cognitive behavioral therapy with large language models: Comparative study with Socratic questioning. *arXiv preprint arXiv:2401.15966*.

<https://arxiv.org/abs/2401.15966>

古川直裕・吉永京子（2024）『責任あるAIとツール』一般社団法人 金融財政事情研究会.

Praxeotech Publishing. (2025). *The AI therapist: How technology is changing the way we heal* (2025 ed.). Artificial Intelligence for Therapy.