

ELIZAから生成AIへ

— ケア対話における利用者体験の比較 —



Livie's Jump, メタデータ株式会社 松田 圭子

対話AIによるケアでは、言語生成の能力だけでなく、話をどう受け止め、どう問いかけるかという対話設計が利用体験を左右すると考えられる。本研究では、同一のLLM（GPT-5.1）上にELIZAを模した反射的対話、汎用LLMによる対話、ケア志向に設計した「やわらかネコ」の3条件を構成し、25名が3条件と対話して評価した。ELIZA型は「理解された」「受け止められた」感覚が他の2条件より有意に低かった。汎用LLMとやわらかネコは同水準だったが、自由記述では前者に「長い・情報が多い」、後者に「優しい／やや回りくどい」との指摘が集まった。受容の「過」と「不足」の両側から、ケア対話の要件を考察する。

背景・目的

対話AIを相談やケアの入口に用いる試みが広がっている。生成AIは高い言語能力を持つが、利用者が「話してよかった」と感じるかどうかは、生成能力そのものより、**話をどう受け止め、どう問い返すかという対話設計**に左右されると考えられる。

1966年のELIZA（Weizenbaum）は、利用者の言葉を反射し短く問い返すだけの単純な仕組みでありながら、多くの人に「理解された」感覚を与えたとされる。では現代の利用者に、この方式は通用するのか。汎用LLMの饒舌な応答や、ケア向けに調整した応答はどう受け取られるのか。

本研究の目的は、反射的対話・汎用LLM・ケア志向設計の3条件を同一のLLM上で比較し、**ケア対話に求められる「過不足のない受容」の要件**を利用者体験から探ることである。



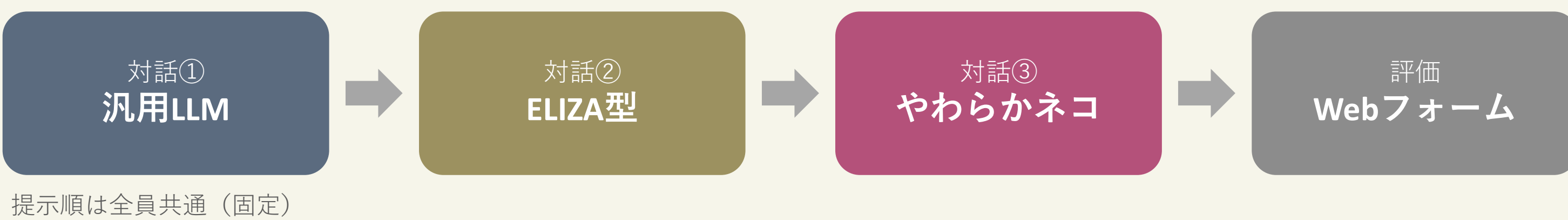
方法

参加者 著者の知人25名（主に学生・社会人、回答期間 2026年10月9日～13日）

手順 各参加者が3条件のAIと順に対話し、Webフォームで評価した。提示順は全員 汎用LLM→ELIZA型→やわらかネコ で固定した

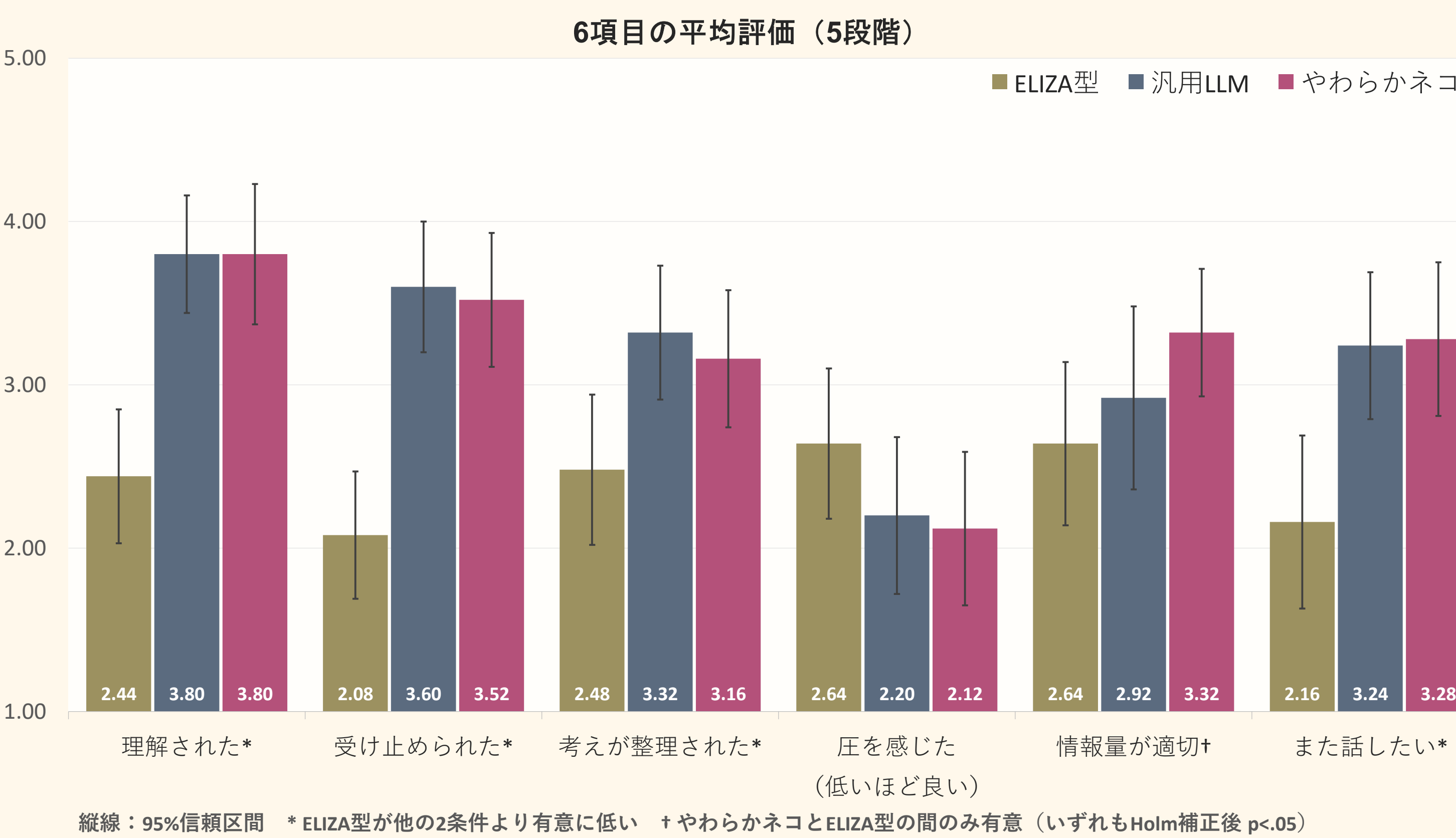
評価 6項目の5段階評価（1=まったくそう思わない～5=とてもそう思う）、自由記述、普段の生成AI利用頻度

分析 項目ごとにFriedman検定（効果量 Kendall's W）を行い、事後比較にWilcoxonの符号付順位検定（Holm補正）を用いた。自由記述は内容により分類した（分類者1名）



提示順は全員共通（固定）

結果：評価



・ELIZA型は「理解された」「受け止められた」「考えが整理された」「また話したい」の4項目で、他の2条件より有意に低かった（Friedman $p \leq .006$, $W = .20 \sim .46$ ）

・汎用LLMとやわらかネコの間には、6項目とも有意差はなかった

・「情報量が適切」はやわらかネコが最も高かった（3.32）が、有意差はELIZA型との間のみ($p = .048$)。「圧を感じた」は3条件間で有意差がなかった（ $p = .085$ ）

・ELIZA型だけ評価が大きく割れた。右表参照→
「また話したい」は「考えが整理された」と強く連動し（ $p = .79$ ）、
「圧を感じた」とはほぼ無関係だった（ $p = .15$ ）。
問い返しで考えが整理できた人には、反射的対話も働いていた。

Q3\Q6	1	2	3	4	5
1	5				
2	5	3			1
3		4	1	1	
4			3		1
5				1	

ELIZA型：Q3 考えの整理 × Q6 また話したい（人数）

限界と今後

・参加者は著者の知人（主に学生・社会人）で、やわらかネコが本来想定する利用者、たとえば相談相手を見つけにくいワンオペ育児・介護の担い手とは、ニーズが異なる可能性がある

・25名の小規模な試行で、話題は参加者ごとに異なる。提示順を汎用LLM→ELIZA型→やわらかネコで固定したため、順序・慣れ・疲労の影響を排除できていない

・汎用LLMには長さや書式を指定しておらず、「長い」という評価には書式指定の差も影響しうる。「話しやすさ」「認知的負荷」は代替の設問で測った

今後の開発課題：ユーザのコンディションや望み（話題の重さ、答えが欲しいのか聞いてほしいのか）を推定し、接し方と方針をゆるやかに切り替える実装を検討する。ユーザが明示的に選択する実装もあってよいのでは。あわせて、想定利用者層での検証、提示順のカウンターバランス、長さや書式をそろえた汎用条件（選択肢）の追加、尺度の精緻化を進める。

謝辞 本研究の実施にあたり、メタデータ株式会社 代表取締役社長 野村直之博士より助言とご支援をいただいた。また、「やわらかネコ」をご利用いただいているユーザの皆様、評価実験にご協力いただいた皆様に深く感謝いたします。

参考文献 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.

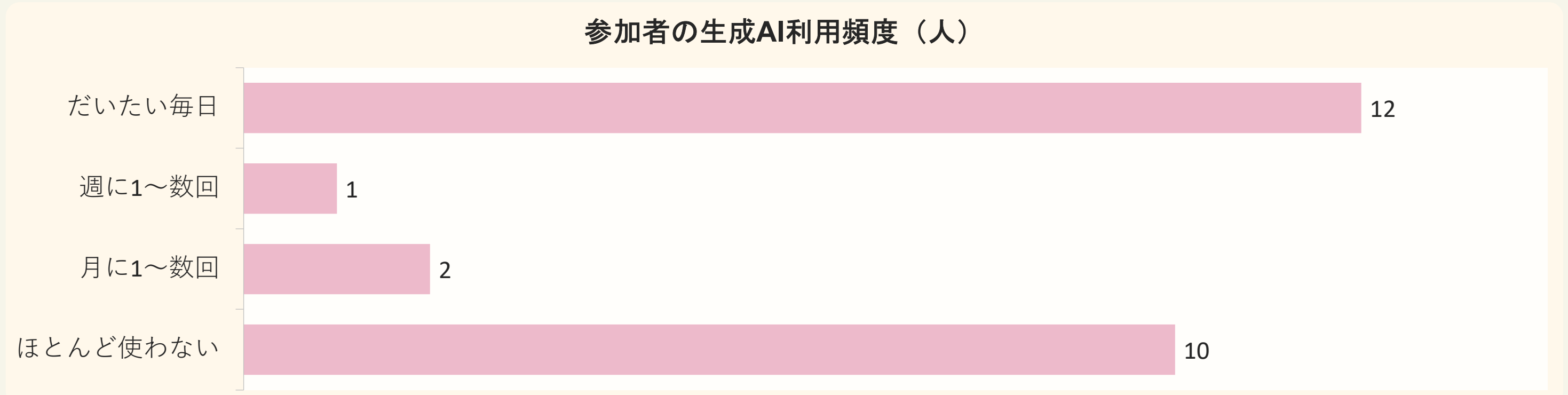
比較した3条件

条件	設計方針とプロンプトの要点
ELIZA型 (提示2番目)	1960年代ELIZAのDOCTORモードを模倣。反射・言い換え・短い問い返しで会話を続ける 分析・原因の推測・助言・励まし・要約・箇条書きはしない。1回の応答は1～2文、質問は1つ
汎用LLM (提示1番目)	特別な調整なし。「自然な日本語で対話し、相談や話題に適切に応答する」という2文の指示のみ
やわらかネコ (提示3番目)	ケア志向に設計した対話AI。心理カウンセリングの基礎、選択理論、認知療法などの知識をもとに、傾聴・受容し「そっとはげます」方針（詳細は非公開）

3条件とも同じLLM（GPT-5.1）を使用し、違いはシステムプロンプトのみとした。

GPT-5.1は、感情の把握と論理的な応答のバランスを考慮して選定した。

設問	本研究での観点
Q1 話を理解してもらえた	理解された感覚
Q2 気持ちや状況を受け止めてもらえた	受容感
Q3 話すことで考えが整理された	自己思考の促進
Q4 AIの発言に圧を感じた	問いかけの圧迫感（低いほど良い）
Q5 応答の情報量は適切だった	認知的負荷（代替指標）
Q6 またこのAIと話したい	話しやすさ・継続意向（代替指標）



結果：自由記述と考察

条件	肯定的な記述	否定的な記述
ELIZA型	短くて快適・気持ちに気づける・もっと話したい（5）	問い返しが面倒・尋問のよう・堂々めぐり（7） 理解されていない・ズレ・表層的（5） 話が進まない（2）
汎用LLM	結論が明快・実用的・安心感（6）	長い・情報が多い・Markdownがそのまま（7）
やわらかネコ	優しい・親しみやすい・選択肢がある（4）	回りくどい・表現が大げさ・続けにくい（4）

自由記述19件。（ ）は件数で、1件が複数の条件に触れる場合がある。AIを名指ししていない2件は内容からELIZA型に分類した

参加者が求めていたもの	やわらかネコとのずれ（自由記述より）
聞いたことに、まず短く答える	寄り添いを優先して「少し回りくどい」
軽い話題は、軽く受け止める	普通の職場付き合いに「自分を守った行動」など、受け止めが重い
「もう大丈夫」と示したら深追いしない	問い返しが続き「3番目も（ELIZA型の質問返しに）ちかい」

受容の姿勢は評価された（理解・受容は汎用LLMと同水準）が、受け止めの強さが話題の重さに合っていなかった

考察：受容の「不足」と「過」

不足 ELIZA型は、問いに答えずに反射と問い返しを続けたため、「理解されていない」「尋問のよう」と受け取られた。反射だけでは、現代の利用者に「理解された」感覚は生まれにくい。一方、自由記述では「尋問」が多く挙がったのにQ4（圧）では差が出ず、負担は「圧」より「面倒さ・疲れ」として経験された可能性がある。

過 汎用LLMは理解・受容の評価は高いが、長さや助言の多さが負担になっていた。やわらかネコは理解・受容を同水準に保ちつつ情報量の評価がやや良かったが、「回りくどい」「大げさ」という、受け止めの言葉の過剰が残った。

過不足のない受容の要件（仮説）

- 問いに答えたうえで、問い返しは控えめが望ましいか（対話の継続への工夫が必要）
- 応答は短く、書式は会話に合わせる
- 「もう大丈夫」と示されたら掘り下げをやめる
- 受け止めの言葉の強さを話題の重さに合わせる