Exempt Research Information Sheet

STUDY TITLE: Expert Perspectives on How AI Chatbots Should Respond to Interpersonal Conflicts (Online Annotation Study)

FOR QUESTIONS ABOUT THE STUDY, CONTACT: Stephane Hatgis-Kessell, stephhk@stanford.edu, 917-972-0943

DESCRIPTION: You are invited to participate in a research study on how AI chatbots, such as ChatGPT or Claude, should respond when people ask for advice about everyday conflicts with friends, partners, family members, or coworkers. We are inviting people with graduate training in mental health counseling to share their professional judgment.

You will be asked to answer a few questions to confirm that you are eligible and about your professional background. You will then read short guidelines describing one counseling approach (for example, cognitive behavioral therapy), with brief comprehension quizzes along the way. Finally, you will read a series of short, hypothetical scenarios, each describing a person's everyday conflict with someone else, and indicate what an ideal response guided by that counseling approach would look like, by selecting from options and writing brief explanations. The scenarios were written with the help of an AI model, do not describe real people, and do not involve crisis situations such as self-harm or abuse. You will not be asked about your own relationships or your clients. We will use your answers to test whether AI models can make similar judgments.

Please do not include information that could identify you, your clients, or anyone else in your written answers.

If you do not pass the comprehension quizzes, the study may end early. You will be compensated for the time you spent in line with Prolific's policies.

Participation in this study is voluntary, and you are free to withdraw your consent at any time.

TIME INVOLVEMENT: Your participation will take approximately 1 hour, in a single online session.

RISKS AND BENEFITS: The risks associated with this study are no more than minimal. Some scenarios describe interpersonal conflicts that may feel familiar, which could cause mild discomfort; you may skip any question or stop at any time. Another risk is a breach of confidentiality, which we minimize as described below. You will not receive any direct benefit from this study, but your answers may help improve how AI chatbots respond to people seeking advice.

PAYMENTS: You will receive $25 through Prolific as payment for completing the study. If the study ends early, you will be compensated for the time you spent in line with Prolific's policies.

PRIVACY AND CONFIDENTIALITY: We will collect your Prolific ID only to process your payment and to prevent duplicate participation. We will not collect your name or contact information. After payments are complete, your Prolific ID will be replaced with a random study

Exempt Research Information Sheet

ID. Study data will be stored securely on Stanford University systems, and only the research team will have access to it.

As part of our analysis, your de-identified answers may be given to AI models, including models run by third-party AI providers. Your Prolific ID is never sent to these providers.

Your answers will be kept confidential. We have procedures to protect your privacy, and your identity will not be included in written materials resulting from the study.

PARTICIPANT'S RIGHTS: If you have read this form and have decided to participate in this project, please understand your participation is voluntary and you have the right to withdraw your consent or discontinue participation at any time without penalty or loss of benefits to which you are otherwise entitled. The alternative is not to participate. You have the right to refuse to answer particular questions. You may withdraw by closing the study window or returning the study on Prolific.

The results of this research study may be presented at scientific or professional meetings or published in scientific journals. Your individual privacy will be maintained in all published and written data resulting from the study.

CONTACT INFORMATION:

Questions: If you have any questions, concerns or complaints about this study, its procedures, risks and benefits, contact the Protocol Director, Stephane Hatgis-Kessell, at stephhk@stanford.edu or 917-972-0943.

INDEPENDENT CONTACT: If you are not satisfied with how this study is being conducted, or if you have any concerns, complaints, or general questions about the study or your rights as a participant, contact the Stanford Institutional Review Board (IRB) to speak to someone independent of the study team at 650-723-2480 or toll free at 1-866-680-2906. You can also write to the Stanford IRB at irbnonmed@stanford.edu.

Save or print a copy of this page for your records.

If you agree to participate in this study, click "I agree" below to begin.