

Taking *STEPS* Forward: Enhancing Online Peer-Counseling with Schema Therapy via Socratic Questioning

Beng Heng Ang¹, Sujatha Das Gollapalli², See-Kiong Ng²

¹Integrative Sciences and Engineering Programme, National University of Singapore

²Institute of Data Science, National University of Singapore

bengheng.ang@u.nus.edu, {idssdg,seekiong}@nus.edu.sg

Abstract

Peer counseling is essential in online mental health communities to provide relatable support to those seeking help, but the peer counselors often lack professional training in therapeutic counseling to produce the desired cognitive changes. In this paper, we present *STEPS*, an AI-powered assistive dialog tool for peer counseling. Unlike other existing tools, *STEPS* assists peer counselors in facilitating *cognitive change* in online counseling settings. Towards this goal, we emulate two key phases in a Schema Therapy-based in-person counseling session—(1) *Schema Assessment* to uncover the deep-seated irrational beliefs underlying an individual’s mental health problems, and (2) *Cognitive Change* to reframe these beliefs into healthier alternatives. In both phases, we employ Socratic questioning techniques to effectively elicit critical introspection and guide cognitive change. We describe *STEPS* and present expert evaluation studies on its counseling conversations on real-world mental health forum posts. Our results indicate that *STEPS* significantly outperforms competitive baselines on all key metrics related to schema assessment, cognitive change strategies, and critical thinking, achieving an impressive average rating of 5 out of 6, highlighting its strong potential as a transformative tool for online peer counseling.

1 Introduction

Online mental health communities (*OMHCs*), such as 7-Cups,¹ Beyond Blue,² and Patient,³ have emerged as cost-effective, scalable, and accessible avenues for those seeking mental health support amid the rising global mental health crisis [Yao *et al.*, 2022; United Nations, 2015]. Peer counseling is essential in online mental health communities by providing relatable support and offering a safe space to share experiences and coping strategies. Through providing empathetic listening and emotional support [Salsabila *et al.*, 2020;

Syed *et al.*, 2024; Ali *et al.*, 2015], as well as sharing their own experiences and practical advice, the peer counselors are able to create a sense of connection with the support seekers [Wang *et al.*, 2012; Joo *et al.*, 2022].

However, to achieve lasting and positive behavioral and cognitive changes in the concerned individuals, deeper insights into their emotional and cognitive needs [Beck, 2020; Young *et al.*, 2006] will be required. Unfortunately, online peer counselors often lack the extensive professional training in therapeutic counseling techniques required to effectively address these needs. As such, AI-powered tools such as *CARE* [Hsu *et al.*, 2023] and *Partner* [Sharma *et al.*, 2021], are now in use to help peer counselors engage with support-seekers, focusing on using motivational interviewing and empathy-based responses. We go beyond these basic approaches by asking: *How can we design a therapy-based assistive tool to help the peer counselors effectively elicit cognitive and behavioral change in online support-seekers?*

We developed *STEPS* (Schema Therapy-Enhanced Peer-Counseling through Socratic Questioning), an AI-powered dialog assistant to address precisely this question. *STEPS* harnesses the generative capabilities of state-of-the-art large language models (*LLMs*) to emulate two key phases of Schema Therapy (*ST*)-based in-person counseling sessions, viz, (1) *Schema Assessment* where deep-seated maladaptive beliefs contributing to one’s mental health problems are uncovered, and (2) *Cognitive Change* where these beliefs are reframed into healthier alternatives. Central to exploring these phases are Socratic questions—specific types of open-ended questions that encourage critical examination about one’s beliefs and perspectives [Paul and Elder, 2019; Montazeri, 2022; Ang *et al.*, 2023; Waltman *et al.*, 2020]. Table 1 illustrates a *STEPS*-generated conversation snippet, highlighting various Socratic questions and their types (highlighted and labeled) in the corresponding counseling phases.

Section 2 summarizes the key components of *STEPS*, followed by an expert evaluation by licensed *ST* practitioners in Section 3. Supplementary details on implementation, LLM prompts, experiments, and evaluation are available in the appendices⁴ along with a demonstration video.⁵

¹<https://www.7cups.com/community/>

²<https://forums.beyondblue.org.au/>

³<https://patient.info/forums>

⁴Appendices are available in <https://github.com/NUS-IDS/ems-mentalhealth/tree/IJCAI.STEPS>

⁵<https://youtu.be/htY8CwqXzr0>

SCHEMA ASSESSMENT

(Opening Post) **Client:** Why do I have low self-esteem and lack confidence?
I always need alcohol to feel better and use that as an excuse.

...

STEPS: I can sense how challenging it must be to feel that way.
What thoughts went through your mind as you are questioning your self-esteem and confidence? [Questions probing for beliefs]

...

STEPS: Do you often feel embarrassed around other people because you think you don't measure up to them in terms of your accomplishments? [Clarifying Questions on the identified EMS (Failure)]

COGNITIVE CHANGE

Client: I often feel embarrassed because I didn't finish college like some of my friends, and I don't have a high-paying job. It feels like I'm always falling short

...

STEPS: I understand how hard it must be to constantly compare yourself to your friends...
What about focusing on the fact that your unique contributions are valued by others, making your accomplishments meaningful... [Questions suggesting healthier alternative beliefs]

Table 1: Conversation snippet in a **STEPS** session. Socratic questions (in blue) are used to encourage critical introspection of beliefs.

2 Methods

Schema Therapy (*ST*) is an integrative therapy framework developed by Young [Young *et al.*, 2006] for treating personality disorders and other mental health problems. *ST*-based counseling has been increasingly adopted for its effectiveness in addressing individual and relationship problems [Masley *et al.*, 2012; Bakos *et al.*, 2015; Taylor *et al.*, 2017]. A key concept in *ST* is Early Maladaptive Schemas (*EMS*), negative and enduring cognitive patterns and beliefs formed in childhood experiences and affect one's perspectives, emotions, and behaviors [Young *et al.*, 2006]. For example, individuals with the "Failure" schema often believe that they are just incompetent.⁶ Identifying these *EMS* is central to *Schema Assessment* [Young *et al.*, 2006]. Effective counseling support challenges these *EMS* and reframes them using structured techniques such as Socratic questioning to foster critical introspection and *Cognitive Change* [Young *et al.*, 2006; Paul and Elder, 2019]. In **STEPS**, these phases are:

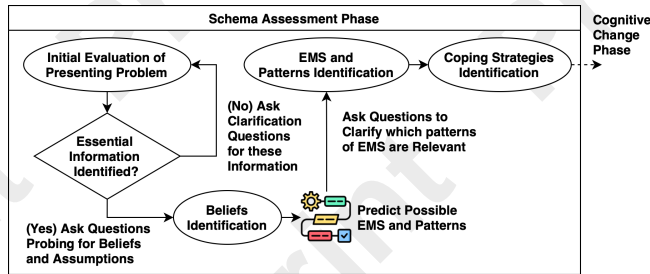


Figure 1: Schema Therapy framework incorporated in **STEPS** is shown partially with the complete workflow and Socratic Questions included in Appendix D.

1. Schema Assessment that focuses on understanding the support seeker's problem by asking relevant questions that gather essential details on *Schema Triggers* (i.e. situations

⁶All 18 *EMS* types are described in Appendix A. See footnote #4

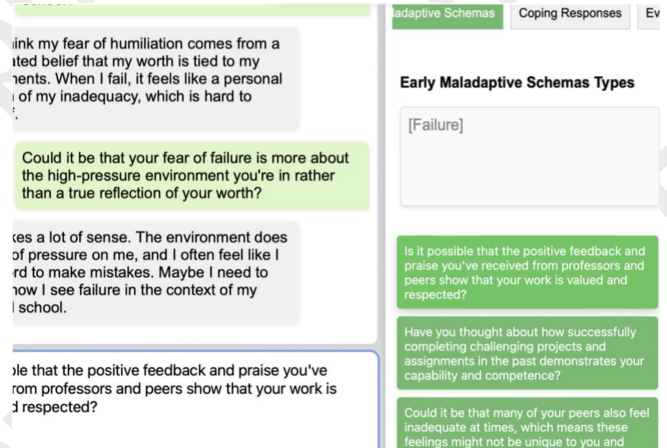


Figure 2: Socratic questions suggested for peer counselors.

that activate negative beliefs), *Thoughts, Emotions, Behaviors, and Bodily Responses* (e.g. "When you start to feel really bored, what does that feel like?"). The support seeker's beliefs are then elicited through probing questions from the problem (e.g. "What does *this* say about you?"). Next, *EMS* are identified from the support seeker's problem and stated beliefs. We extend a recent *EMS* prediction study [Gollapalli *et al.*, 2023], which compared support seeker's beliefs with the Young Schema Questionnaire (*YSQ*) [Young and Brown, 2003], a psychometric tool of 232 statements on *EMS*, using various textual similarity, *LLMs*, and textual entailment-based techniques. Building on this approach, we compare the sentence similarity between the beliefs of the support seekers and the *YSQ* statements, selecting those *EMS* with scores above a threshold⁷ for further validation through clarifying questions (e.g. the second question in Table 1).

⁷Sentence transformers [Reimers and Gurevych, 2019] are used with a similarity threshold of $\theta = 0.4$

Configuration	Assessment	Cognitive	Inquiring	Imagining	Doing	Reflecting
<i>LLM_{ST+phases}</i>	1.58 ± 1.44	1.16 ± 1.40	1.67 ± 1.61	1.75 ± 1.76	1.50 ± 1.24	1.25 ± 1.22
STEPS	5.08 ± 0.67	5.08 ± 0.51	4.83 ± 0.71	5.00 ± 0.60	5.00 ± 0.85	5.33 ± 0.49

Table 2: Mean ratings and standard deviations for *STEPS* and *LLM_{ST+phases}* across all evaluation metrics.

2. Cognitive Change focuses on encouraging support seekers to critically evaluate their *EMS* and explore more balanced, healthier beliefs through questions that probe for alternative strategies, (e.g. “What other ways could you manage this belief?”), evidence-based inquiry (e.g. “What past experiences support your belief that you are a failure?”), and challenging their beliefs through alternative perspectives (e.g. “Could it be that other students also struggled with the test, and this doesn’t mean you are a failure?”).

Empathetic Socratic Question Generation: Socratic questioning is central to *STEPS*, where empathetic responses combined with specific open-ended questions are systematically employed throughout the counseling phases. *STEPS* uses recent *LLMs* to generate these Socratic questions. While *LLM* can generate empathetic responses and context-specific Socratic questions in isolated (one-step) contexts [Li *et al.*, 2022; Priya *et al.*, 2023; Ang *et al.*, 2023], research has shown that their planning capabilities are insufficient for steering conversations in goal-oriented contexts such as negotiation or counseling. To address this shortcoming, a “dialog policy” is often used to guide the *LLM* on the appropriate response type [Jang *et al.*, 2022; Deng *et al.*, 2024; Gollapalli and Ng, 2025]. Given our counseling objectives and lack of training data, we directly incorporate the “dialog policy” suggested by *ST* framework to determine specific Socratic questions to generate within each phase. The precise prompts provided in Appendix C are used to generate the specific types of Socratic questions presented in Figure 1. *STEPS* is designed as an assistive tool in a human-in-the-loop framework, presenting suggestions for further editing before the final interaction with the support seeker. Part of our system interface illustrating this aspect is shown in Figure 2 with the full interface detailed in Appendix D.⁴

3 Experiments and Results

Experimental Setup: We adopt the recent approach for conversation evaluation where two separate *LLMs* role-play the conversation actors [Abbasiantaeb *et al.*, 2024]. That is, *STEPS* acts as a “peer counselor” and responds to *LLM_{Patient}* simulating a “support seeker”. We used 20 CounselChat forum posts [Bertagnoli, 2020] available from a prior study on *ST* [Gollapalli *et al.*, 2023] and evaluate two configurations: (1) *LLM_{ST+phases}*, a *LLM* prompted to emulate a Schema therapist by incorporating **Schema Assessment** and **Cognitive Change** into its dialog and (2) *STEPS*, our proposed system. Our goal is to test if the existing SOTA *LLMs* with parameterized knowledge of *ST* are able to emulate the process designed to enable cognitive change in *ST* counseling given appropriate prompts, compared to *STEPS*.⁸

⁸We used the “gpt-4o-2024-05-13” model in experiments but based on initial comparisons, expect the results to apply to all *LLMs*

Evaluation Criteria and Metrics: We evaluate *LLM_{ST+phases}* and *STEPS* on their generated conversations along specific metrics of (1) *Schema Exploration and Assessment (Assessment)*, and (2) *Application of Cognitive Change Techniques (Cognitive)* by collecting human ratings on a scale from 0 to 6. These metrics were established for evaluating a Schema therapist’s proficiency in identifying *EMS* and applying cognitive strategies to reframe *EMS* into healthier beliefs [Young and Fosse, 2005] (Appendix E). Additionally, we evaluate if the conversations foster critical thinking, a key aspect of effective counseling [Overholser and Beale, 2023], by adapting metrics from the Organisation for Economic Cooperation and Development critical thinking rubrics [Vincent-Lancrin *et al.*, 2019]. These metrics evaluate proficiency in: (3) *Assumption-checking (Inquiring)*, (4) *Identifying alternatives and hypotheticals (Imagining)*, (5) *Justifying reasoning with logical premises (Doing)*, and (6) *Assessing limitations of a perspective (Reflecting)*.⁹ Following prior work [Gollapalli *et al.*, 2023], we hired two certified Schema therapists and one Psychology graduate with experience in *ST* via Upwork.⁹ They evaluated 30 generated dialogs at a rate of five samples per hour, with compensation of \$20–\$40 per hour. Six dialogs were rated independently by all evaluators, yielding an inter-rater agreement of 0.72 ($p < .05$) as measured by Kendall’s τ [Kendall, 1938], indicating strong consensus among evaluators.

As shown in Table 2, *STEPS* achieved an overall average rating of ~ 5 , corresponding to a “very good” rating, and significantly outperformed the *LLM_{ST+phases}* by 3–4 points. This highlights *STEPS*’s potential for enhancing online peer counseling via *ST*. Furthermore, the low variability in ratings, with standard deviations of ~ 0.6 compared to the baseline’s ~ 1.4 , highlights the more consistent performance in *STEPS* compared to that of *LLM_{ST+phases}*.

4 Conclusion

This paper presents *STEPS*, an AI-powered dialog assistant that integrates Schema Therapy principles and Socratic questioning with *LLMs* to enhance online peer counseling. Expert evaluation by counseling professionals shows that *STEPS* significantly outperforms the *LLM* baseline in identifying and addressing *EMS* and fostering critical thinking. However, in high-stakes domains like mental health support, it is critical that there are additional layers of oversight to ensure that the presented responses are safe, appropriate, and therapeutically sound. Before deployment in *OMHCs*, these findings must be carefully validated in real-world peer counseling environments. Additional safeguards, such as a controller to filter for safe *LLM* responses from *STEPS*, should also be in place.

with similar capabilities. Our prompts are included in Appendix C.

⁹www.upwork.com

Ethics Statement

This research strictly adheres to the ACM Code of Ethics, emphasizing the ethical implications of AI usage in mental health contexts. *STEPS* is developed as an AI-assistive tool designed to support peer counselors in facilitating introspective conversations and promoting positive change, rather than serving as a replacement for professional therapy. The system operates within a human-in-the-loop framework, where all AI-generated suggestions from *STEPS* are reviewed and mediated by trained peer counselors. To ensure a thorough evaluation of *STEPS*, we engaged accredited Schema therapists through Upwork, offering fair compensation for their time (details provided in Section 3). The general ethical and privacy concerns for use of AI in health care are not applicable for our prototype system that has not been tested/deployed in real settings and all datasets are based on publicly-available mental health forum posts.

Acknowledgments

This research/project is supported by the National Research Foundation, Singapore under its AI Singapore Programme (AISG Award No: AISG-GC-2019-001-2B) and the National University of Singapore, Institute of Data Science. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not reflect the views of the National University of Singapore.

References

- [Abbasiantaeb *et al.*, 2024] Zahra Abbasiantaeb, Yifei Yuan, Evangelos Kanoulas, and Mohammad Aliannejadi. Let the llms talk: Simulating human-to-human conversational qa via zero-shot llm-to-llm interactions. In *Proceedings of the 17th ACM International Conference on Web Search and Data Mining*, pages 8–17, 2024.
- [Ali *et al.*, 2015] Kathina Ali, Louise Farrer, Amelia Gulliver, Kathleen M Griffiths, et al. Online peer-to-peer support for young people with mental health problems: a systematic review. *JMIR mental health*, 2(2):e4418, 2015.
- [Ang *et al.*, 2023] Beng Heng Ang, Sujatha Das Gollapalli, and See Kiong Ng. Socratic question generation: A novel dataset, models, and evaluation. In *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics*, pages 147–165, 2023.
- [Bakos *et al.*, 2015] Daniela Schneider Bakos, Alex Eduardo Gallo, and Ricardo Wainer. Systematic review of the clinical effectiveness of schema therapy. *Contemp Behav Health Care*, 1(1):11–15, 2015.
- [Beck, 2020] Judith S Beck. *Cognitive Behavior Therapy: Basics and Beyond*. Guilford Press, 2020.
- [Bertagnolli, 2020] Nicolas Bertagnolli. Counsel chat: Bootstrapping high-quality therapy data. <https://github.com/nbertagnolli/counsel-chat>, 2020.
- [Deng *et al.*, 2024] Yang Deng, Wenxuan Zhang, Wai Lam, See-Kiong Ng, and Tat-Seng Chua. Plug-and-play policy planner for large language model powered dialogue agents. In *The Twelfth International Conference on Learning Representations*, 2024.
- [Gollapalli and Ng, 2025] Sujatha Das Gollapalli and See-Kiong Ng. Pirsuader: A persuasive chatbot for mitigating psychological insulin resistance in type-2 diabetic patients. In *COLING*, pages 5997–6013, 2025.
- [Gollapalli *et al.*, 2023] Sujatha Das Gollapalli, Beng Heng Ang, and See Kiong Ng. Identifying {Early Maladaptive Schemas} from mental health question texts. In *The 2023 Conference on Empirical Methods in Natural Language Processing*, 2023.
- [Hsu *et al.*, 2023] Shang-Ling Hsu, Raj Sanjay Shah, Prathik Senthil, Zahra Ashktorab, Casey Dugan, Werner Geyer, and Diyi Yang. Helping the helper: Supporting peer counselors via AI-empowered practice and feedback. *arXiv preprint arXiv:2305.08982*, 2023.
- [Jang *et al.*, 2022] Youngsoo Jang, Jongmin Lee, and Kee-Eung Kim. GPT-critic: Offline reinforcement learning for end-to-end task-oriented dialogue systems. In *International Conference on Learning Representations*, 2022.
- [Joo *et al.*, 2022] Jin Hui Joo, Lee Bone, Joan Forte, Erin Kirley, Thomas Lynch, and Hanan Aboumatar. The benefits and challenges of established peer support programmes for patients, informal caregivers, and healthcare providers. *Family practice*, 39(5):903–912, 2022.
- [Kendall, 1938] Maurice G Kendall. A new measure of rank correlation. *Biometrika*, 30(1-2):81–93, 1938.
- [Li *et al.*, 2022] Qintong Li, Piji Li, Zhaochun Ren, Pengjie Ren, and Zhumin Chen. Knowledge bridging for empathetic dialogue generation. In *Proceedings of the AAAI conference on artificial intelligence*, volume 36, pages 10993–11001, 2022.
- [Masley *et al.*, 2012] Samantha A Masley, David T Gillanders, Susan G Simpson, and Morag A Taylor. A systematic review of the evidence base for schema therapy. *Cognitive behaviour therapy*, 41(3):185–202, 2012.
- [Montazeri, 2022] Mohammad Sadegh Montazeri. *Psychotherapist’s Guide to Socratic Dialogue*. Springer, 2022.
- [Overholser and Beale, 2023] James C Overholser and Eleanor Beale. Socratic questioning and guided discovery. *Psychotherapy Skills and Methods That Work*, page 74, 2023.
- [Paul and Elder, 2019] Richard Paul and Linda Elder. *The thinker’s guide to Socratic questioning*. Rowman & Littlefield, 2019.
- [Priya *et al.*, 2023] Priyanshu Priya, Kshitij Mishra, Palak Totala, and Asif Ekbal. Partner: A persuasive mental health and legal counselling dialogue system for women and children crime victims. In *IJCAI*, pages 6183–6191, 2023.
- [Reimers and Gurevych, 2019] Nils Reimers and Iryna Gurevych. Sentence-bert: Sentence embeddings using siamese bert-networks. In *Proceedings of the 2019*

Conference on Empirical Methods in Natural Language Processing. Association for Computational Linguistics, 11 2019.

- [Salsabila *et al.*, 2020] Salsabila Salsabila, Jessica Wiryan-tara, Nurnadiah Salsabila, and Muhammad Afif Alhad. The role of peer counseling on mental health. *Bisma The Journal of Counseling*, 4(3):242–253, 2020.
- [Sharma *et al.*, 2021] Ashish Sharma, Inna W Lin, Adam S Miner, David C Atkins, and Tim Althoff. Towards facilitating empathic conversations in online mental health support: A reinforcement learning approach. In *Proceedings of the Web Conference 2021*, pages 194–205, 2021.
- [Syed *et al.*, 2024] Sara Syed, Zainab Iftikhar, Amy Wei Xiao, and Jeff Huang. Machine and human understanding of empathy in online peer support: A cognitive behavioral approach. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*, pages 1–13, 2024.
- [Taylor *et al.*, 2017] Christopher DJ Taylor, Penny Bee, and Gillian Haddock. Does schema therapy change schemas and symptoms? a systematic review across mental health disorders. *Psychology and Psychotherapy: Theory, Research and Practice*, 90(3):456–479, 2017.
- [United Nations, 2015] United Nations. Transforming our world: The 2030 agenda for sustainable development. 2015. *View Article*, 2015.
- [Vincent-Lancrin *et al.*, 2019] Stéphan Vincent-Lancrin, Carlos González-Sancho, Mathias Bouckaert, Federico De Luca, Meritxell Fernández-Barrerra, Gwénaél Jacotin, Joaquin Urgel, and Quentin Vidal. *Fostering Students’ Creativity and Critical Thinking: What It Means in School*. *Educational Research and Innovation*. ERIC, 2019.
- [Waltman *et al.*, 2020] Scott H Waltman, Lynn M McFarr, and Bret A Moore. *Socratic Questioning for Therapists and Counselors*. Routledge, 2020.
- [Wang *et al.*, 2012] Yi-Chia Wang, Robert Kraut, and John M Levine. To stay or leave? the relationship of emotional and informational support to commitment in on-line health support groups. In *Proceedings of the ACM 2012 conference on computer supported cooperative work*, pages 833–842, 2012.
- [Yao *et al.*, 2022] Zheng Yao, Haiyi Zhu, and Robert E Kraut. Learning to become a volunteer counselor: Lessons from a peer-to-peer mental health community. *Proceedings of the ACM on Human-Computer Interaction*, 6(CSCW2):1–24, 2022.
- [Young and Brown, 2003] Jeffrey Young and G Brown. Ysq-13. *New York, NY: Schema Therapy Institute*, 2003.
- [Young and Fosse, 2005] J Young and G Fosse. Schema therapy rating scale for individual therapy sessions (strs-i-1). *Schema Therapy Institute*, 2005.
- [Young *et al.*, 2006] Jeffrey E Young, Janet S Klosko, and Marjorie E Weishaar. *Schema Therapy: A Practitioner’s Guide*. Guilford Press, 2006.