

CoSim: LLM-based Simulation System Trains Student Academia and Life Supporters to Communicate with College Students Experiencing Stress and Anxiety

Jie Cai
Tsinghua University
Beijing, China
jie-cai@mail.tsinghua.edu.cn

Yuxuan Li
Carnegie Mellon University
Pittsburgh, Pennsylvania, USA
yuxuanll@andrew.cmu.edu

Wenjing Deng
Tsinghua University
Beijing, China
dengwj24@mails.tsinghua.edu.cn

Chun Yu
Tsinghua University
Beijing, China
chunyu@mail.tsinghua.edu.cn

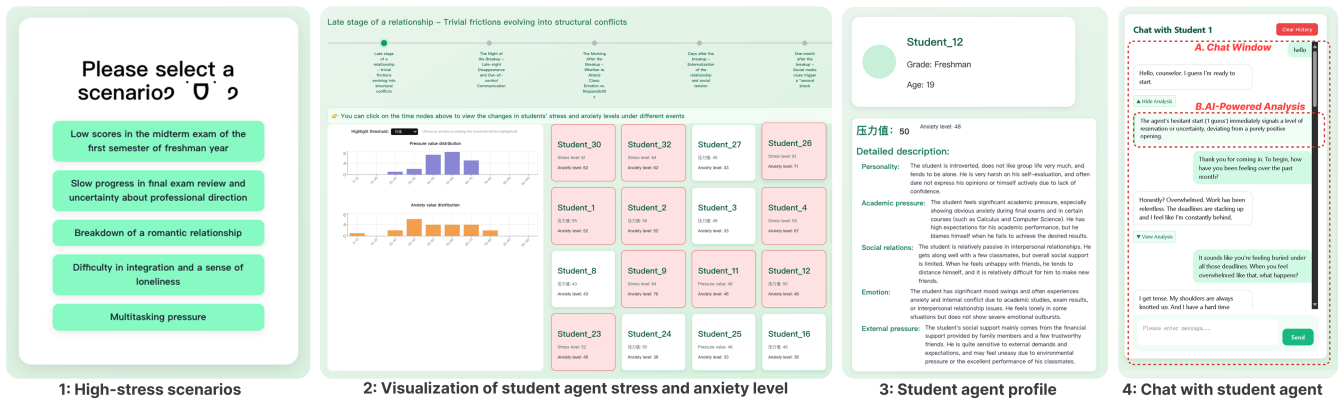


Figure 1: (1) SALSs select a scenario relevant to their practice and common challenges. (2) They review an overview of student agents’ stress and anxiety levels (red indicates high stress and anxiety). (3) They pick a student agent with high stress and anxiety and inspect profile and background. (4) They chat with the student agent using each reply and the corresponding inner monologue.

Abstract

Student Academia and Life Supporters (SALSs) at Chinese universities are a unique group often taking multiple roles (e.g., tutors, advisors, and psychological counselors) for student administration, but lack socio-technical support when dealing with students with stress and anxiety. LLM provides the opportunity to run simulations and provide rehearsal opportunity. Through a formative study involving five SALSs and 33 students, we developed CoSim, an LLM-based simulation system featuring 33 built-in agents that simulate these students’ stress and anxiety in diverse scenarios, achieving 72% and 76% similarity. The think-aloud evaluation with 10 SALSs shows that (1) CoSim resonates with real students in expression, reasoning, and trust, with some limitations; (2) CoSim supports SALSs’ reflection of communication strategies, but also causes an extra burden. This study contributes to the intersection of health and education through LLM-based simulation, offering

design opportunities to support diverse stakeholders in education settings.

CCS Concepts

• **Human-centered computing** → **Empirical studies in collaborative and social computing**; **Empirical studies in HCI**.

Keywords

LLM, Mental Health, College Student, Stress and Anxiety, Simulation, Education Technology

1 Introduction

College students often experience stress and anxiety arising from academic demands, social transitions, and limited access to counseling resources [1, 14, 22]. HCI research has devoted attention and carried out work on student mental health, including self-tracking tools [7, 15, 32, 35], conversational agents [19], and reflection or mindfulness applications [2, 10, 16]. While these systems provide valuable support, they tend to emphasize reactive interventions, engaging students when distress becomes visible or self-reported.

This paper was prepared for PoliSim@CHI 2026: LLM Agent Simulation for Policy, a workshop at CHI 2026 (CHI Conference on Human Factors in Computing Systems), April 16, 2026, Barcelona, Spain.

Yet, a proactive response is critical for preventing problems from escalating [11, 27, 33]. In practice, educators and counselors often play a critical role in noticing subtle early warning signs through everyday interactions with students. However, they face several challenges, such as limited training in mental health, fragmented information about students' experiences, and difficulties in identifying underlying causes of stress [9, 13]. Despite their critical role, relatively few HCI efforts have focused on supporting these stakeholders in developing skills and tools for proactive intervention.

In China, these needs are particularly acute, with a large student population and a shortage of professional mental health practitioners [4, 9, 23]. As a result, much of the responsibility for frontline mental support falls to a unique role known as Fudaoyuan [21], commonly translated as Student Academia and Life Supporters (SALSs). They engage closely with students and are expected to detect early signs of psychological well-being issues and intervene when necessary. However, they typically lack formal mental health training [9]. This gap motivates exploration of approaches that provide opportunities for SALSs or other stakeholders to engage in structured, low-risk practice.

Recent advances in large language models (LLMs) have enabled the development of social agents capable of simulating plausible human behaviors, emotions, and interactions [25, 26]. Within education, such agents have been applied to classroom simulation [18, 29, 36], teachable student agents [12, 34], and scenario-based tutor training [6, 8, 17, 24, 31], demonstrating their potential to provide realistic yet controllable practice environments. In contrast, their use for student mental health contexts remains underexplored, particularly preventive training for non-specialist counselors such as SALSs.

To better understand the pressing needs of SALSs, we conducted a formative study with five SALSs. The interviews focused on their daily practices and cases of students experiencing stress or anxiety. Analysis revealed three recurring challenges: difficulties in timely identification of at-risk students, limited access to longitudinal information about students' trajectories, and uncertainty in tracing problems back to their underlying causes. Building on these findings, we introduce CoSim, an LLM-based simulation system that generates student agents grounded in interview and survey data, each exhibiting plausible stress and anxiety responses across common university scenarios. CoSim provides a multi-phase interface that enables counselors to (1) monitor groups of students, (2) explore individual profiles and behavioral trajectories, and (3) engage in interactive counseling with real-time feedback on stress, anxiety, and conversational effectiveness.

To evaluate the system, we conducted a user study with 10 SALSs, guided by the following research questions:

- RQ1: How do SALSs feel about the simulation system design?
- RQ2: How do SALSs feel about the interaction with the simulated students?
- RQ3: How do the simulated students and conversations influence SALSs' communication and self-reflection?

2 Study Design

We first interviewed five SALSs to understand their support practices and challenges and to collect cases and scenarios that evoke

student stress and anxiety. We then deployed a structured interview chatbot with a prescribed stress-and-anxiety protocol and used the resulting transcripts to construct student agents. Next, we developed CoSim by combining these scenarios with the student agents to simulate student stress and anxiety, invited 10 SALSs to use CoSim for about an hour, and analyzed the resulting dialogues to identify SALSs' communication patterns.

2.1 Formative Study to Identify Challenges

We interviewed five SALSs to explore the challenges they face in providing student support. Through thematic analysis, we identified three challenges: (C1) Talking to stressed students is challenging for beginners. SALSs currently lack low-risk opportunities to rehearse supportive conversations before stepping into high-stakes offline interactions. (C2) SALSs often struggle to access a student's background or behavioral history. This lack of integrated information makes it difficult to pinpoint the exact sources of a student's stress during a conversation. (C3) SALSs often worry about missing subtle conversational cues. Furthermore, they often remain uncertain whether their verbal responses support the student or unintentionally cause further stress. These challenges underscore the need for a simulation-based context that enables SALSs to rehearse before similar talks with real peers.

2.2 Design and Implementation of CoSim

To address these challenges, we developed CoSim with three integrated features. First, we developed five typical scenarios that can cause student stress and anxiety, such as low exam scores or heavy multitasking pressure. This approach adopts scenario-based design [28]. The student agents are constructed using an interview-grounded method [26]. We interviewed 32 junior students using a semi-structured protocol focused on academic and life stress. We then used these demographics and interview transcripts to construct student agents for each student. Once an student agent has gone through each timestamp in the scenarios, it will complete a stress survey [5, 30] and receive a normalized score indicating its stress level.

After selecting a scenario and an student agent, SALSs can inspect the student agent's stress level and profile, which provides a full view of the student agent's background before the chat starts. Finally, SALSs chat with the student agent in real time and read the agent's inner monologue to see the direct impact of their words on the student agent's feelings. We also add a distrust prompt that elicits more authentic, sometimes negative reactions, thereby simulating realistic conversations (students may withhold or deflect when talking with SALSs) and mitigating LLM "always reply" tendencies.

2.3 User Study

We recruited 10 SALSs (4 females, 6 males) to use CoSim for about an hour, with diverse backgrounds through university forums and social media. On average, participants had 2.5 years of experience working as SALSs. To better capture their usage behaviors of our system, we chose remote participation over on-site to screen-record everything about their interaction with our system for follow-up data analysis. Meanwhile, we used think-aloud with remote screen

sharing to watch and record everything. The think-aloud method allowed users to verbalize their thoughts and internal processes while performing a task to gain insight into their cognitive experiences, problem-solving strategies, and decision-making [3, 20]. We first briefly introduced our system with screensharing to demo the interface-specific features and functions. Next, we asked the participants to open our system and share their screen with their consent to record the whole session. Third, we especially asked the participant to think about a scenario they usually deal with students or one they would like to deal with students. Finally, we guided them to chat with a stressed and anxious student to understand their situation and provide suggestions to help the student as much as possible. The think-aloud session lasted about 40-50 minutes, and after the session, we conducted a 20-30-minute interview to ask their general feeling about the system, their thoughts about the simulated students, and their experiences with communication and reflection, which are tied closely with our research questions.

3 Preliminary Findings

We report three key findings. First, the simulated agents achieved a 76% alignment score with their real-student counterparts on mental health questionnaires, indicating strong behavioral fidelity. Second, SALSs reported that interacting with the simulated agents helped them better understand diverse student experiences and practice communication strategies. Third, the simulation enabled SALSs to observe psychological dynamics that would be difficult or risky to reproduce in real settings, thereby supporting more informed and anticipatory reflective communication.

References

- [1] Randy P. Auerbach, Philippe Mortier, Ronny Bruffaerts, Jordi Alonso, Corina Benjet, Pim Cuijpers, Koen Demuyttenaere, David D. Ebert, Jennifer Greif Green, Penelope Hasking, Elaine Murray, Matthew K. Nock, Stephanie Pinder-Amaker, Nancy A. Sampson, Dan J. Stein, Gemma Vilagut, Alan M. Zaslavsky, Ronald C. Kessler, and WHO WMH-ICS Collaborators. 2018. WHO World Mental Health Surveys International College Student Project: Prevalence and distribution of mental disorders. *Journal of Abnormal Psychology* 127, 7 (Oct. 2018), 623–638. doi:10.1037/abn0000362
- [2] Aishwarya Chandrasekaran, London Bielickie, Diya Shah, Harisha Janakiraman, and Matthew Louis Mauriello. 2025. "I spent 14 hours debugging just one assignment": Toward Computer-Mediated Personal Informatics for Computer Science Student Mental Health. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. ACM, Yokohama Japan, 1–19. doi:10.1145/3706598.3713269
- [3] Elizabeth Charters. 2003. The Use of Think-aloud Methods in Qualitative Research An Introduction to Think-aloud Methods. *Brook Education Journal* 12, 2 (July 2003). doi:10.26522/brocked.v12i2.38
- [4] Yumeng Chen, Yali Zhang, Guoliang Yu, and . 2022. Prevalence of mental health problems among college students in mainland China from 2010 to 2020: A meta-analysis. *Advances in Psychological Science* 30, 5 (May 2022), 991–1004. doi:10.3724/SP.J.1042.2022.00991
- [5] Sheldon Cohen, Tom Kamarck, and Robin Mermelstein. 1983. A Global Measure of Perceived Stress. *Journal of Health and Social Behavior* 24, 4 (1983), 385–396. doi:10.2307/2136404
- [6] Wei Dai, Yi-Shan Tsai, Jionghao Lin, Ahmad Aldino, Hua Jin, Tongguang Li, Dragan Gašević, and Guanliang Chen. 2024. Assessing the proficiency of large language models in automatic feedback generation: An evaluation study. 7 (2024), 100299. doi:10.1016/j.caeai.2024.100299
- [7] Alyssa J Gatto, Yasuo Miyazaki, and Lee D Cooper. 2022. Help me help myself: examining an electronic mental health self-monitoring system in college students. *Higher Education* 83, 1 (2022), 163–182. doi:10.1007/s10734-020-00646-8
- [8] Dollaya Hirunyasiri, Danielle R. Thomas, Jionghao Lin, Kenneth R. Koedinger, and Vincent Alevan. 2023. Comparative Analysis of GPT-4 and Human Graders in Evaluating Praise Given to Students in Synthetic Dialogues. arXiv:2307.02018 [cs] doi:10.48550/arXiv.2307.02018
- [9] Rui Hou, SiLang Huang, Kenneth Po-Lun Fung, Alan Li, Cunxian Jia, Shengli Cheng, Jianguo Gao, Jingxuan Zhang, and Josephine Pui-Hing Wong. 2024. Who is helping students? A qualitative analysis of task-shifting and on-campus mental health services in China's university settings. *Social Science & Medicine* 363 (Dec. 2024), 117527. doi:10.1016/j.socscimed.2024.117527
- [10] Jennifer Huberty, Jeni Green, Christine Glissmann, Linda Larkey, Megan Puzia, and Chong Lee. 2019. Efficacy of the Mindfulness Meditation Mobile App "Calm" to Reduce Stress Among College Students: Randomized Controlled Trial. *JMIR mHealth and uHealth* 7, 6 (June 2019), e14273. doi:10.2196/14273
- [11] Zamira Hyseni Duraku, Holly Davis, and Era Hamiti. 2023. Mental health, study skills, social support, and barriers to seeking psychological help among university students: a call for mental health support in higher education. 11 (2023), 1220614. doi:10.3389/fpubh.2023.1220614
- [12] Hyoungwook Jin, Minju Yoo, Jeongeon Park, Yokyung Lee, Xu Wang, and Juho Kim. 2025. TeachTune: Reviewing Pedagogical Agents Against Diverse Student Profiles with Simulated Students. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. Association for Computing Machinery, New York, NY, USA, 1–28. doi:10.1145/3706598.3714054
- [13] Shantalea Johns. 2024. Exploring Faculty and Academic Advisors' Experiences Supporting College Students with Mental Health Challenges: A Qualitative Study Using Interpretative Phenomenological Analysis. 5, 2 (2024), 63–73. doi:10.12930/NACR-23-03
- [14] Payton J. Jones, So Yeon Park, and G. Tyler Lefevor. 2018. Contemporary College Student Anxiety: The Role of Academic Distress, Financial Stress, and Support. *Journal of College Counseling* 21, 3 (Oct. 2018), 252–264. doi:10.1002/jocc.12107
- [15] Taewan Kim, Haesoo Kim, Ha Yeon Lee, Hwarang Goh, Shakhboz Abdigaporov, Mingon Jeong, Hyunsung Cho, Kyungsik Han, Youngtae Noh, Sung-Ju Lee, and Hwajung Hong. 2022. Prediction for Retrospection: Integrating Algorithmic Stress Prediction into Personal Informatics Systems for College Students' Mental Health. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (New Orleans, LA, USA) (CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 279, 20 pages. doi:10.1145/3491102.3517701
- [16] Heera Lee and Andrea Kleinsmith. 2019. Public Speaking Anxiety in a Real Classroom: Towards Developing a Reflection System. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, Glasgow Scotland UK, 1–6. doi:10.1145/3290607.3312875
- [17] Jionghao Lin, Zifei Han, Danielle R. Thomas, Ashish Gurung, Shivang Gupta, Vincent Alevan, and Kenneth R. Koedinger. 2024. How Can I Get It Right? Using GPT to Rephrase Incorrect Trainee Responses. arXiv:2405.00970 [cs] doi:10.48550/arXiv.2405.00970
- [18] Yiping Ma, Shiyu Hu, Xuchen Li, Yipei Wang, Yuqing Chen, Shiqing Liu, and Kang Hao Cheong. 2025. When LLMs Learn to be Students: The SOEI Framework for Modeling and Evaluating Virtual Student Agents in Educational Interaction. arXiv:2410.15701 [cs] doi:10.48550/arXiv.2410.15701
- [19] Raju Maharjan, Darius Adam Rohani, Per Bækgaard, Jakob Bardram, and Kevin Doherty. 2021. Can we talk? Design Implications for the Questionnaire-Driven Self-Report of Health and Wellbeing via Conversational Agent. In *Proceedings of the 3rd Conference on Conversational User Interfaces (CUI '21)*. Association for Computing Machinery, New York, NY, USA, 1–11. doi:10.1145/3469595.3469600
- [20] Julia M. Markel, Steven G. Opferman, James A. Landay, and Chris Piech. 2023. GPTEach: Interactive TA Training with GPT-based Students. In *Proceedings of the Tenth ACM Conference on Learning @ Scale (L@S '23)*. Association for Computing Machinery, New York, NY, USA, 226–236. doi:10.1145/3573051.3593393
- [21] MEPRC Ministry of Education of People's Republic Of China. 2017. *Regulations on the Construction of Counsellors in Colleges and Universities, Decree No.43 of the Ministry of Education of the People's Republic of China.*. https://www.gov.cn/gongbao/content/2017/content_5244874.htm
- [22] Mohammad Mofatteh. 2020. Risk factors associated with stress, anxiety, and depression among university undergraduate students. *AIMS Public Health* 8, 1 (2020), 36–65. doi:10.3934/publichealth.2021004
- [23] Xuan Ning, Josephine Pui-Hing Wong, Silang Huang, Yina Fu, Xiaojie Gong, Lizeng Zhang, Carla Hilario, Kenneth Po-Lun Fung, Miao Yu, Maurice Kwong-Lai Poon, Shengli Cheng, Jianguo Gao, and Cun-Xian Jia. 2022. Chinese University Students' Perspectives on Help-Seeking and Mental Health Counseling. 19, 14 (2022), 8259. doi:10.3390/ijerph19148259 Publisher: Multidisciplinary Digital Publishing Institute.
- [24] Sitong Pan, Robin Schmucker, Bernardo Garcia Bulle Bueno, Salome Aguilar Llanes, Fernanda Albo Alarcón, Hangxiao Zhu, Adam Teo, and Meng Xia. 2025. TutorUp: What If Your Students Were Simulated? Training Tutors to Address Engagement Challenges in Online Learning. *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (2025)*. doi:10.48550/arXiv.2502.16178 arXiv:2502.16178 [cs].
- [25] Joon Sung Park, Joseph O'Brien, Carrie Jun Cai, Meredith Ringel Morris, Percy Liang, and Michael S. Bernstein. 2023. Generative agents: interactive simulacra of human behavior. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*. ACM, San Francisco CA USA, 1–22. doi:10.1145/3586183.3606763

- [26] Joon Sung Park, Carolyn Q. Zou, Aaron Shaw, Benjamin Mako Hill, C. Cai, Meredith Ringel Morris, Robb Willer, Percy Liang, and Michael S. Bernstein. 2024. Generative Agent Simulations of 1,000 People. *ArXiv abs/2411.10109* (2024). doi:10.48550/arXiv.2411.10109 arXiv:2411.10109 [cs].
- [27] Nicola Reavley and Anthony F. Jorm. 2010. Prevention and early intervention to improve mental health in higher education students: a review. 4, 2 (2010), 132–142. doi:10.1111/j.1751-7893.2010.00167.x
- [28] Mary Beth Rosson and John M. Carroll. 2007. SCENARIO-BASED DESIGN. In *The Human-Computer Interaction Handbook* (2 ed.). CRC Press. Num Pages: 20.
- [29] Robin Schmucker, Meng Xia, Amos Azaria, and Tom Mitchell. 2023. Ruffle&Riley: Towards the Automated Induction of Conversational Tutoring Systems. arXiv:2310.01420 [cs] doi:10.48550/arXiv.2310.01420
- [30] Robert L. Spitzer, Kurt Kroenke, Janet B. W. Williams, and Bernd Löwe. 2006. A Brief Measure for Assessing Generalized Anxiety Disorder: The GAD-7. *Archives of Internal Medicine* 166, 10 (May 2006), 1092–1097. doi:10.1001/archinte.166.10.1092
- [31] John Stamper, Ruiwei Xiao, and Xinying Hou. 2024. Enhancing LLM-Based Feedback: Insights from Intelligent Tutoring Systems and the Learning Sciences. arXiv:2405.04645 [cs] doi:10.48550/arXiv.2405.04645
- [32] Rui Wang, Weichen Wang, Alex daSilva, Jeremy F. Huckins, William M. Kelley, Todd F. Heatherston, and Andrew T. Campbell. 2018. Tracking Depression Dynamics in College Students Using Mobile Phone and Wearable Sensing. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 2, 1, Article 43 (March 2018), 26 pages. doi:10.1145/3191775
- [33] Christian J. Wiedermann, Verena Barbieri, Barbara Plagg, Pasqualina Marino, Giuliano Piccoliori, and Adolf Engl. 2023. Fortifying the Foundations: A Comprehensive Approach to Enhancing Mental Health Support in Educational Policies Amidst Crises. 11, 10 (2023), 1423. doi:10.3390/healthcare11101423
- [34] Songlin Xu, Hao-Ning Wen, Hongyi Pan, Dallas Dominguez, Dongyin Hu, and Xinyu Zhang. 2025. Classroom Simulacra: Building Contextual Student Generative Agents in Online Education for Learning Behavioral Simulation. doi:10.1145/3706598.3713773 arXiv:2502.02780 [cs].
- [35] Jingxin Yu, Amid Ayobi, Paul Marshall, and Aisling Ann O’Kane. 2025. Creatively Supporting Mental Wellbeing: A Tangible Toolkit to Scaffold Self-Tracking through Mindful Colouring. In *Proceedings of the Nineteenth International Conference on Tangible, Embedded, and Embodied Interaction (TEI ’25)*. Association for Computing Machinery, New York, NY, USA, Article 23, 17 pages. doi:10.1145/3689050.3704944
- [36] Zheyuan Zhang, Daniel Zhang-Li, Jifan Yu, Linlu Gong, Jinchang Zhou, Zhanxin Hao, Jianxiao Jiang, Jie Cao, Huiqin Liu, Zhiyuan Liu, Lei Hou, and Juanzi Li. 2025. Simulating Classroom Education with LLM-Empowered Agents. In *Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers)*, Luis Chiruzzo, Alan Ritter, and Lu Wang (Eds.). Association for Computational Linguistics, Albuquerque, New Mexico, 10364–10379. doi:10.18653/v1/2025.naacl-long.520