Practicing Stress Relief for the Everyday: Designing Social Simulation Using VR, AR, and LLMs

ANONYMOUS AUTHOR(S)

Fig. 1. We show participants XR and LLM-powered prototypes that simulate everyday stressors to understand perspectives on simulation for self-care. Pictured are a virtual reality simulation with breathwork guidance for practicing a social party (left) as well as an augmented reality scene for practicing public speaking (right).

Stress is an inevitable part of day-to-day life yet many find themselves unable to manage it themselves, particularly when professional or peer support are not always readily available. As self-care becomes increasingly vital for mental well-being, this paper explores the potential of social simulation as a safe, virtual environment for practicing stress relief for everyday situations. Leveraging the immersive capabilities of VR, AR, and LLMs, we developed eight interactive prototypes for various everyday stressful scenarios (e.g. public speaking) then conducted prototype-driven semi-structured interviews with 19 participants. We reveal that people currently lack effective means to support themselves through everyday stress and found that social simulation fills a gap for simulating real environments for training mental health practices. We outline key considerations for future development of simulation for self-care, including risks of trauma from hyper-realism, distrust of LLM-recommended timing for mental health recommendations, and the value of accessibility for self-care interventions.

CCS Concepts: • **Human-centered computing → Empirical studies in HCI**.

Additional Key Words and Phrases: virtual reality, large language models, mental health, design, speed dating

ACM Reference Format:

Anonymous Author(s). 2025. Practicing Stress Relief for the Everyday: Designing Social Simulation Using VR, AR, and LLMs. In *Proceedings of ... (CHI '25)*. ACM, New York, NY, USA, 30 pages. https://doi.org/XXXXXXX.XXXXXXX

Manuscript submitted to ACM

1 INTRODUCTION

When faced with stress and anxiety, taking care of ourselves is challenging. At least 70% of Americans experience physical and mental symptoms of stress and 58% of people aged 18 to 34 feel most days contain an overwhelming amount of stress. However, only 37% of people believe they manage this stress effectively [2]. Although mental health resources have become more accessible through advancements like virtual therapy and online communities [16, 43], most people do not have constant access to these resources when going about their personal, work, or social lives and inevitably encounter stress or even panic. Stress is not only difficult to feel, but also affects our ability to fulfill other aspects of our lives as family members, friends, professionals, and community members [59, 67].

Human-Computer Interaction (HCI) research has extensively explored mental health care needs and systems over the past few decades [5]. Despite the many people who are not able to manage their own stress effectively, though, there continues to be a notable gap in applying HCI technology to practicing everyday self-care. One potential avenue for offering people a safe environment to practice their stress relief skills is with *simulation technology* – technological systems that aim to replicate real-world behaviors, agents, or processes in a virtual environment [38, 70, 93]. Virtual simulation can aid in stress-relief practices by providing immersive, controlled scenarios that allow individuals to safely engage with and manage stressors. By recreating realistic social and environmental dynamics, users can potentially practice coping mechanisms in a low-risk environment, a preventative self-care measure that can improve emotional regulation and preparedness for real-life situations [13].

In our work, we leverage virtual reality (VR), augmented reality (AR), and large language models (LLMs) to develop social simulation in order to simulate common everyday environments for users to practice self-soothing – specifically, box breathing, which is a technique well-documented in literature for its effectiveness in treating stress, anxiety, and panic [18]. Using eight low fidelity prototypes, we conducted *prototype-driven interviews* with 19 participants to explore their reactions to using simulation for practicing stress relief for the everyday. Specifically, our work explores the research question:

RQ: What are the needs, risks, and design considerations in using simulation technology to help practice management of everyday stress?

Our findings indicate that many people lack effective methods for self-care and found our simulation prototypes useful for building skills to deal with everyday stressors; most expressed they would incorporate this use of simulation into their real-world practices as a self-care practice tool. We also studied different dimensions for designing social simulation for self-care, including modality, interactivity, and mental health guidance features. Participants particularly preferred AR due to its integration with real surroundings, which made it easier to transfer skills to real-world contexts and engage physically with the environment. However, participants also saw accessibility as a valuable asset of LLM-powered, purely text-based simulation despite lower immersion, given its likely greater accessibility on-the-go. Our interviews also uncovered trade-offs in designing social simulation for self-care, such as strong preferences for more realism alongside concerns about potential trauma from hyper-realistic designs. Finally, our work highlights key design considerations for using LLMs in social simulation, such as users' concerns about the effects of LLM-generated recommendations on their sense of agency and real-world transferability of skills.

2 RELATED WORK

In this section, we review relevant literature to our work, including HCI for mental health, simulation technology, extended reality for mental health, and large language models for mental health.

2.1 HCI for Mental Health and Self-Care

HCI work for mental and emotional health has spanned decades of research, ranging from computational approaches involving machine learning and natural language processing [5, 96], to more qualitative understandings of people's needs in digital mental health [52, 73, 97], to development of technical HCI systems involving wearables and VR interventions [25, 32, 37]. There is immense potential (and proven success) that technology for mental health care can have. However, most work in HCI scholarship has focused on how to deliver support from a professional therapeutic standpoint, such as internet-based cognitive-behavioral therapy and dialectic therapy [42, 63, 80], or online peer and social support [3, 86, 105]. Our work fills a different need in support, focusing on helping individuals practice skills to *self*-care, particularly during their day-to-day encounters.

Apart from extended reality and mindfulness, which we review below, research in HCI that has studied how people self-care or the efficacy of self-care tools has primarily centered on tracking and self-management for chronic, diagnosable conditions. For example, work has studied how people navigate and manage their own chronic depression through reliance on social ties [12] and built systems for people to track their lifestyle habits for bipolar disorder [6]. Additionally, self-care research has primarily centered on *physical health* conditions. For example, a review paper in 2015 that studied self-care technologies in HCI [65] regarding chronic conditions reviewed 29 papers in HCI, where just three included any mental health conditions (bipolar disorder and unipolar depression), while a review done in 2020 on self-care measures for adolescents reviewed health needs for diabetes, cystic fibrosis, and female care without any mental health papers included. Regardless, some aspects of this past literature may still be applicable to self-care for the everyday stress and anxiety people experience, such as the need for accessibility of self-management tools [65].

2.2 Simulation in HCI

Agent-based simulation has long been used in the social sciences to explain and test different aspects of human behavior [34, 95]. Simulation in HCI have traditionally focused on virtual behaviors in online networks [1, 72, 81], or used agent-based simulation to make design decisions for online communities [76, 77]. In recent years, accurate human dialogue and complex social dynamics has also been able to be replicated using simulation techniques through the rise of large language models and generative AI [70, 83]; for example, Park et al.'s work was able to use LLM-powered "generative agents" who act as believable human-like avatars in a virtual town engaging in everyday behaviors like forming interpersonal relationships and coordinating plans [69].

One significant advantage of agent-based simulations, compared to direct experimentation, is their ability to identify downstream and unintended effects in a safe, controlled environment [76, 77]. For example, it can be used to test new algorithms safely before deployment to a real-world community, which is particularly important for the potentially vulnerable and sensitive context of health [54]. Since simulations can accurately mirror real-world processes without the associated risks, they provide a safer space for individuals to practice self-care skills at their own pace and in a more comfortable atmosphere.

2.3 Immersive Technology for Mental Health

2.3.1 Virtual and Augmented Reality Technology for Mental Health and Self-Care. Among the wide range of technologies that have been leveraged to aid people's mental health, such as chatbots [4] and wearables [22], extended reality (XR) including both VR and AR has emerged as perhaps the most rapidly growing and promising tool. VR interventions in particular have shown effectiveness in helping people develop sensory awareness and flow [11, 82], conduct mindfulness and meditation practices [41, 98, 106], and to treat a range of mental health conditions [5, 8, 29, 31, 61, 66, 75].

Apart from treatment for particular conditions, HCI and XR research has also studied mental health interventions for the everyday, novice user. This is perhaps most apparent in HCI's focus on meditation and mindfulness through XR [101]; researchers have used VR and AR to deliver learning curriculuum to teach meditation skills [30], practice varying breathing techniques [74, 100], help pedestrians navigate more mindfully [17], and designed multimodal experiences to facilitate walking meditation [94]. XR also allows valuable embodiment of emotions and thoughts into physical space, allowing users to "physically" interact such as punching away negative thoughts [35]. Given that our work focuses on environmental and social simulation of scenarios like public speaking or social anxiety, we also drew from past work studying the effectiveness of XR interventions for accurately simulating anxiety (even for non-anxious people) [60] and its ability to effectively reduce stress for experiences like public speaking [36, 53]. However, we note a distinction in our work from past XR systems for mindfulness that have removed people from real-world environments altogether, directing them to calming settings like abstract environments [64], natural environments [14], or otherworldly spaces [62]. Overall, XR for self-care has been a rich field of study in HCI, and shown immense promise in its ability to improve symptoms and help people more tangibly interact with their own mental and emotional states.

2.3.2 Large Language Models (LLMs) for Mental Health. There has been remarkable success and growth with applying large language models to the mental health domain whether in the clinical setting or non-professional context. LLMs have been used to directly provide therapeutic help [33, 51, 55], classify mental health experiences [103], and teach social and coping skills [40, 83, 104]. There is also emerging work on the capabilities of LLMs to help people gain mental health insights for themselves. For example, LLMs can help patients with their journaling content and then summarize patients' thoughts for their caretakers [45] and help suggest restructuring of negative thoughts [85].

LLMs have been used to power conversational agents and mental health chatbots for both care-seekers and care-providers. LLM-powered agents have shown they are able to provide context-aware responses that can be helpful for providing mental health support directly to people, such as through providing direct treatment [44] and helping reframe thought processes [58]. LLM-based chatbots can also offer guidance to support-providers such as helping them provide better responses and in diagnosis [15, 84]. However, there are also dangers of LLMs being used in digital mental health intervention, including delivering of misinformation and exhibiting bias [24, 107]. Given these risks, we focus on moving forward in building simulation for self-care in human-centered ways through first understanding human needs, preferences, and boundaries when encountering these technologies. This study probes users' thoughts on LLM-powered guidance, but does not use LLMs to provide direct mental health guidance advice; instead, we restrict LLMs to producing realistic conversational dialogue [57] and gather user reactions to LLM-recommended timing for delivering breathing guidance instead.

Our work lies at the intersection of the above fields. Specifically, we adopt a simulation-based approach, leveraging XR and LLMs to explore users' opinions on being placed in real-world environments to practice self-coping and self-care skills. Our study differentiates itself from prior literature by *applying simulation to teach self-care practices*, and moreover focuses on people's *everyday routine stressors* rather than those with specific conditions.

3 METHODS

In this section, we describe the three key design dimensions - **modality**, **interactivity**, and **guidance needs** (3.1) - that guide our prototype development (3.2). Using these prototypes, we conducted ***prototype-driven interviews*** (3.3) and analyzed these interviews using thematic analysis (3.4).

Our prototype development uses an approach akin to parallel prototyping [27] in order to generate richer insights into people's design preferences and lead to higher quality design recommendations. Given our exploratory goals to uncover users' needs and constraints, we designed eight different prototypes rather than testing one "optimal" solution so we could thoroughly explore the design space. While we did probe participants after each prototype (e.g. *"What are your initial impressions?"*, *"Would you use this?"*), we also focused on asking overarching thoughts and design feedback after participants had finished seeing all eight prototypes (e.g. "*Which of these designs do you prefer, if any, and why?*", "*How realistic did these simulations feel?*").

3.1 Key Design Dimensions

We designed prototypes following three main dimensions guided by our research question and past studies in immersive technology [79]:

Modality: We aimed to investigate how the form of the simulation — whether VR, AR, or text-based — and thus its level of immersion [79] affects users' perceptions of the technology in managing stress.

Interactivity: We also ask what level and types of interaction users desire within these simulations – whether through dialogue, body language, or environmental feedback. Our prototypes range from non-interactive, static environments to fully interactive settings where users engage in real-time, responsive conversations with LLM-powered avatars.

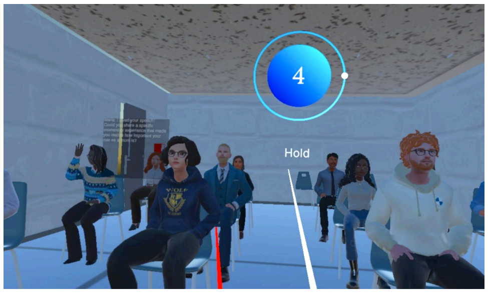
Guidance Needs: Our study explores how to help people practice active soothing mechanisms, such as breathing, during moments of high stress. However, some past research in VR/AR has found success in just facilitating exposure (without any incorporated self-care techniques) to help build resilience [49]. We explore how users felt practicing stress relief techniques superimposed on the simulation of real-world environments; in particular, we selected box breathing as our guidance intervention given its positive effects for stress and relative simplicity to learn for new users [7, 19, 23]. As a result, for each modality we include prototypes without any breathing intervention and those with intervention (**`VR-Guided`**, **`AR-Guided`**, **`Text-Guided`**).

Through these three dimensions, we designed our prototypes with the goal of providing insights into how different aspects of simulation can contribute to social simulation for self-care.

3.2 Design Prototypes

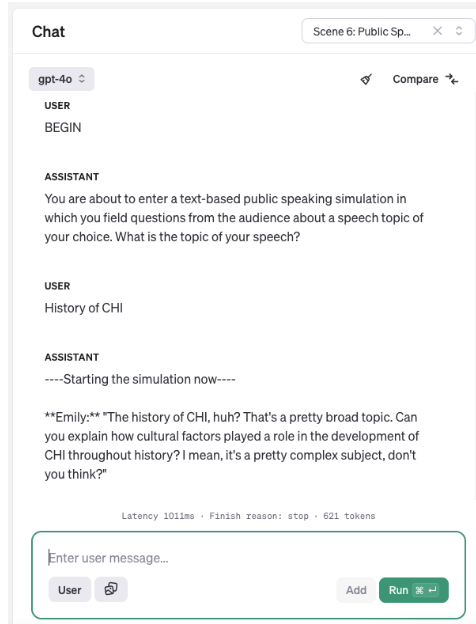
Although everyone experiences everyday stress [46], what triggers that stress varies from person-to-person. To better understand participants' feelings about our prototypes in the context of their own personal experiences, we designed prototypes around **three distinct scenarios** that have been simulated in prior research, with the aim for at least one to resonate with participants as a stressor. Due to interview time constraints given our many prototypes, we did not show participants interactive controls for AR and instead asked participants to imagine given the system they had just previously interacted with in VR. Participants were also shown screenshots of prototypes to better visualize the various scenarios and interfaces.

(1) **Public Speaking Q&A** [89, 92]. Public speaking is the most common fear [28]. Given the practical challenge in asking participants to give a full speech, we opted to imagine themselves giving a speech and interact in a



(a) VR-Guided for public speaking scenario.

(b) AR-Static for public speaking.



(c) Text-Interactive for public speaking.

Fig. 2. Examples of different prototypes, shown here for the public speaking scenario.

post-talk Q&A setting instead. In this scenario, the user is asked to imagine they have just finished giving a speech on any topic of their choosing and answer questions from a 12-person audience.

(2) **Social Party** [39, 68]. We selected to simulate a crowded social situation given that social anxiety is the most common anxiety disorder [91]. In this scenario, the user is asked to imagine they just moved to a new area and have been invited to a party. They are met with a crowded room of people and are greeted in conversation with 4 other party guests.

(3) **Interpersonal Conflict** [83]. We simulated interpersonal conflict (in this case, with a roommate) due to most people having encountered it but lacking the opportunity to practice the complex dynamics of conflict resolution in an emotionally safe setting [26, 83]. Stress during conflict has also been shown to be a barrier to resolution [88]. In this scenario, the user is asked to engage in conflict resolution with their roommate about a topic of their choosing (see Figure 4).

Based on our design dimensions of modality, interactivity, and guidance needs, we created the following eight prototypes for each of the three scenarios. We go into more detail on the implementation in the following sections. Example screenshots are shown in Figure 2.

1 **`VR-Static`**: a non-interactive scene in virtual reality. Avatar(s) and the user do not engage in any dialogue interaction.

2 **`VR-Interactive`**: same as VR-Static, except avatar(s) and the user can engage in dialogue. User can pull up a keyboard to type or dictate their replies.

3 **VR-Guided**: same as VR-Interactive, with the addition that the user can press a button on their controller at any point to pull up a breathing guidance orb. The user can use the breathing guidance at any point for as long as they wish.
4 **AR-Static**: a non-interactive scene.
5 **AR-Interactive**: same as AR-Static, except avatar(s) and the user can engage in dialogue similar to the VR version.
6 **AR-Guided**: same as AR-Interactive, with the addition that the user can press a button on their controller at any point to pull up a breathing guidance orb. The user can use the breathing guidance at any point for as long as they wish.
7 **Text-Interactive**: users interact with a text-based bot that role plays as an audience/party guests/a roommate, depending on the users' scenario.
8 **Text-Guided**: same as Text-Interactive, with the addition of text output of breathing guidance (i.e. "Breathing Guidance: Inhale for 4 seconds. Hold for 4 seconds. Exhale...").

3.2.1 Prototype Development. Our method involves developing usable, interactive prototypes, as opposed to using storyboards, in order to generate more detailed insights as participants could tangibly experience the immersive capabilities of social simulation. We go into more detail below on building the prototypes, including the visual and audio components of VR/AR and the LLM-powered dialogue.

Environment. For all 18 VR/AR builds (three VR and three AR, across three different scenarios) we built the prototypes using the Unity game engine and deployed on a Meta Quest 3 headset. All avatars were created using Ready Player Me[1] where the first three authors selected diverse avatars and appropriate clothing for the scenario (e.g. casual outfits for the social party). Additionally, we looped sound clips that suited each scenario using freely available audio from online. Specifically, we included soft classroom and hallway noises for the public speaking scenario, incoherent talking and dance music for the social party scenario, and layered air conditioning and street noises for the interpersonal conflict scenario.

Dialogue. Dialogue for all prototypes is generated using GPT-4o, the most recent GPT version at the time of this study. Our work contributes to the existing but little work on using LLMs to generate realistic avatar dialogue for interactive extended reality [87, 99]. We used consistent prompting patterns across text-based, VR, and AR prototypes (see Appendix). To ensure natural interactions, the first three authors iterated on prompting patterns to create realistic exchanges. For instance, we asked GPT to generate character personas with seed examples of conversational styles to create realistic and relevant interactions. We also ensured that GPT would output dialogue from individual characters so they spoke one-at-a-time in order to not overwhelm the user. For VR/AR prototypes, we used the OpenAI-Unity package[2] to integrate GPT-4o responses to avatars in VR/AR. For the text-based prototypes, participants interacted directly with GPT-4 via the OpenAI Playground[3]. We conducted pilot testing and gathered feedback from a psychiatry practitioner before proceeding with participant interviews.

Interaction and Controls. The text-based prototypes were created and shown to participants in OpenAI Playground, so participants interacted with GPT directly using a laptop keyboard. For VR/AR interaction, participants interacted with avatars by using the Quest 3 controllers to type on a VR keyboard, or could use speech-to-text dictation that we

[1]https://readyplayer.me/
[2]https://github.com/srcnalt/OpenAI-Unity
[3]https://platform.openai.com/playground

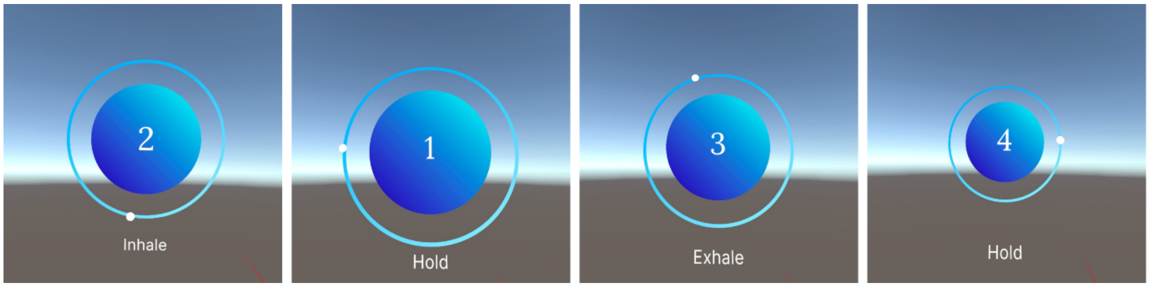


Fig. 3. Onboarding system of box breathing guidance. The user is asked to inhale, hold, exhale, and hold for 4 seconds each. The bubble expands and contracts following the breath, with a countdown timer for 4 seconds for each stage.

Fig. 4. Example of interaction flow in the VR prototype for roommate conflict. The user can select a topic (left) then begin a dialogue with the generated scene (middle). At any point, the user may trigger the breathwork guidance and conduct box breathing while the simulation is paused (right).

implemented using WhisperAI[4]. Participants are stationary in the simulation, although they can turn to look anywhere 360-degrees.

Guidance System Development. For VR/AR prototypes, the box breathing guidance was informed by prior research to use cool colors [56], abstract imagery [94], and expansion and contraction to guide the breathing pattern [94]. Given past work's recommendations [71], we also showed users a dedicated onboarding system (without any scenario simulation) for users to familiarize themselves with box breathing before entering the simulation; screenshots of the box breathing guidance system are shown in Figures 3 and 4. Users could access breathing guidance at any point by pressing a button on the Quest 3 controller. For **`Text-Guided)`**, its guidance implementation is output just as text instructions in OpenAI Playground (see Figure 2); we prompt GPT to output breathwork at recommended moments of high stress. We ask users their opinions on this LLM-powered guidance instruction versus manually-selected guidance (like in VR/AR) in our interviews.

3.3 Semi-Structured Interview

We conducted prototype-driven interviews with 19 participants. Interviews were semi-structured, guided by a list of questions (see Appendix) but allowed to deviate depending on topics participants may have introduced. The interviews took place in-person, lasted approximately 1 hour, and were audio-recorded.

[4]https://openai.com/index/whisper/

Table 1. Participant demographics for gender, age, and race.

Participant	Gender	Race / Ethnicity	Age	Exposure Condition
P1	Cisgender Man	Asian	21	Public Speaking Q&A
P2	Cisgender Woman	Asian	24	Public Speaking Q&A
P3	Cisgender Woman	White	21	Social Party
P4	Cisgender Man	Hispanic, Latino, and/or Spanish Origin	21	Public Speaking Q&A
P5	Cisgender Woman	Black and/or African-American	19	Interpersonal Conflict
P6	Cisgender Man	Asian	30	Public Speaking Q&A
P7	Prefer not to say	Black and/or African-American	29	Interpersonal Conflict
P8	Cisgender Woman	Asian	19	Interpersonal Conflict
P9	Cisgender Man	Black and/or African-American	21	Interpersonal Conflict
P10	Cisgender Man	Asian, White	22	Social Party
P11	Cisgender Man	Asian	42	Social Party
P12	Cisgender Woman	Asian	24	Interpersonal Conflict
P13	Prefer not to say	White	41	Social Party
P14	Cisgender Woman	Asian	29	Social Party
P15	Cisgender Woman	Asian	20	Interpersonal Conflict
P16	Cisgender Man	Hispanic, Latino, and/or Spanish Origin	22	Interpersonal Conflict
P17	Cisgender Man	White	38	Public Speaking Q&A
P18	Cisgender Woman	Black and/or African American	19	Interpersonal Conflict
P19	Cisgender Woman	Asian	23	Public Speaking Q&A

3.3.1 Participants and Recruitment. We aimed to recruit participants of diverse backgrounds, so eligibility was only restricted to those 18 years and above and required in-person presence for the study. We recruited participants through putting up physical flyers in the local community, posting on social media, and snowball sampling. Interested respondents were asked to complete an intake form that asked for their basic demographic information of gender identity, age, and racial identity. Interview participants' demographic information is listed in Table 1. In total, we interviewed 19 participants, who we met in-person. Participants were compensated with a $20 Amazon gift card. This study was approved by the appropriate Institutional Review Board (IRB).

3.3.2 Exposure Condition. Participants were asked to report their expected stress and anxiety levels for the three scenarios: (1) answering questions from an audience after their public speaking, (2) attending a crowded social party, and (3) having an interpersonal conflict with a roommate. Participants rated their stress levels ("On a scale of, how stressful or anxious do you anticipate yourself to be when...") for these three situations using the Subjective Units of Distress Scale (SUDS) scale from 0 to 10, which is an established rating system in psychology [47, 102] (see Appendix for full scale). After participants finished this survey, we recorded participants' highest stress situation and asked participants if they felt comfortable experiencing simulation for it. All participants said yes; we informed participants that at any point they could ask us to change scenarios or stop the interview altogether. For participant safety, we also

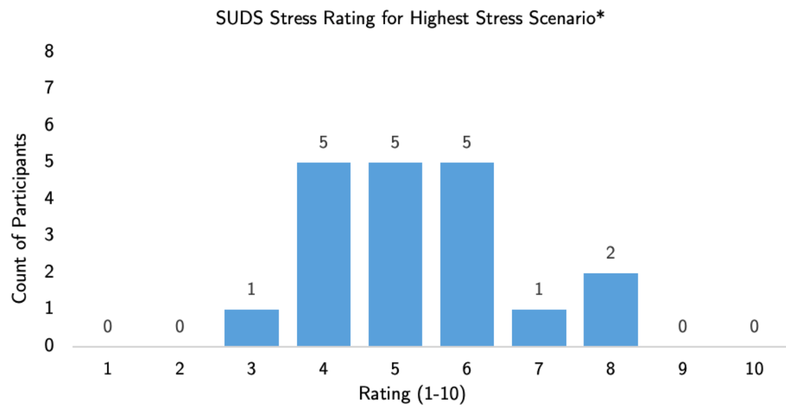


Fig. 5. SUDS score rating among participants for their exposure condition

automatically excluded showing them scenarios for which they rated their anticipated stress 9 or 10, and also provided all participants with the local county mental health hotline number for extra precaution.

3.3.3 Interview Protocol. We showed the eight prototypes on the participant's highest stress topic given their initial survey. In addition, we also showed the "onboarding" guide to the breathing guidance system before the participant saw **VR-Guided** or **AR-Guided** to help users engage with the exercise [71]; in the onboarding guide, the participant saw only the breathing orb and no other simulation (see Figure 3), during which the participant could take as long as they needed to practice the box breathing pattern before proceeding. For each prototype, we provided an introduction blurb for participants to imagine. For example, before showing participants the VR prototypes for interpersonal conflict with a roommate, we stated: "*Imagine that you recently moved in with a new roommate. Things have been tense, though, and you plan to confront your roommate at home later today after work about an ongoing source of disagreement between you two. Nervous about how this conversation will go, you put on a virtual reality simulation to practice what you are going to say to them. In this scene, you'll be met with your roommate who will speak with you about this ongoing conflict. You'll be able to choose the topic of the conflict.*" For the public speaking and interpersonal conflict scenes, participants could choose their own topic of choice for their speaking topic and conflict reason, respectively, or choose from pre-sets as shown in Figure 4.

We aimed to understand people's needs, whether this technology met those needs, and understand the design opportunities to improve self-care technology. The interview protocol consisted of showing each of the eight prototypes and asking participants questions revolving around four main concepts: (1) **initial impressions**, (2) **usefulness** ("*To what extent do you feel this is useful?*", "*How does practicing stress relief in this way compare to your current practices?*") (3) **realism and natural interaction** ("*To what extent did the interactions feel natural or unnatural?*", "*What aspects felt realistic or unrealistic?*") and (4) **general feedback** ("*To what extent did you like this design?*", "*What additional features would be useful?*"). For the prototypes involving the breathing guidance, we additionally asked participants what their thoughts were on using guidance or, if they didn't use guidance, their reasoning for not using the breathing guidance system. Lastly, we asked participants about their experience as a whole with the simulation as well as their general thoughts on applying simulation to their everyday stress relief practices. Given the semi-structured nature of the interviews, our research team's interviewer would also ask follow-ups, clarifications, and ask participants to expand on thoughts they expressed in an evolving nature during the course of the interview. The full list of interview questions are in our appendix.

3.4 Data Analysis

After completing all interviews, we anonymized, transcribed, and coded all 19 interview transcripts. Based on Braun and Clarke's method of thematic analysis [9], our iterative analytic cycle consisted of: (1) recording and transcribing interviews (2) coding transcripts (3) amalgamating codes (4) discussing codes (5) highlighting themes (6) writing and revising memos. Our interviews and analysis continued until saturation, where new interviews gave no new insights. We first generated the lowest-level, granular codes through each transcript, then organized these codes into axial-codes. Afterwards, we arranged these findings into five high-level themes presented in this paper.

In addition to how participants thought of the simulation prototypes in general, we also wanted to measure participants' preferences among the different prototype options (e.g. VR versus AR versus text-based modalities). To produce numerical ratings for the eight prototype concepts, the first two authors independently read participants' reactions to each design concept and rated it from 1 (disliked a lot) to 5 (liked a lot); see Figure 6 for results. Note that these ratings were given by the researchers from analyzing the participants' interviews, rather than the participants directly giving numerical ratings of the prototypes themselves. The authors had high internal consistency with a Cronbach's alpha of 0.88, suggesting that the authors' ratings reliably indicate participants' preferences among the prototypes.

4 RESULTS

We organize our findings around five high-level themes identified through our interview analysis, including people's current experiences of stress (4.1), overall perspectives and experiences using simulation for self-care (4.2), and findings around our three high-level dimensions for designing simulation for self-care: modality (4.3), interactivity (4.4), and guidance (4.5).

4.1 Current Experiences of Everyday Stress

Participants experienced a range of different stressors in their everyday lives, most of which were in the themes of social anxiety and work. 12 out of 19 participants described their everyday experience of stress as related to social anxiety, including being in social settings with people they don't know (P3, P12), being in crowded spaces (P10, P12), giving presentations (P2, P14, P15), and talking on the phone (P19). Other sources of stress include taking exams, meeting deadlines, and generally having a lot of work or personal tasks to complete.

4.1.1 Current methods to handle stress. In addition to asking participants what types of everyday stress they experience, we also asked them how they dealt with stress. While participants did mention a range of coping mechanisms, such as relying on social connections (P11, P14) and problem-solving (P3, P10, P14, P15), the most common theme across participants was using distraction.

> "*talking to people about lighthearted things...so I can **distract** myself*" (P5)
> "*[I use] messaging applications...talk about random stuff and **distract** myself from any conflicts going on.*" (P16)
> "***distract** myself a bit, browsing social media*" (P1)
> "*I usually **distract** myself*" (P12)

"*Not a technique, but I do **distract** with hobbies that are not related*" (P16)

Regarding strategies targeted towards mitigating their own stress, many participants had tried habits like journaling (P14, P17), meditation and meditation apps (P3, P12, P14), and breathing techniques (P1, P2, P4, P5, P11); however, almost all did not sustain these practices over time with the exception of regular exercise, which was cited as effective at keeping chronic stress low (P4, P7, P8, P9, P13). Participants' reasons for not continuing use of self-care techniques during stressful situations ranged from finding some techniques **difficult** ("*Journaling was a thing that my therapist said, but I don't have patience for it. . . I wish it would be more straightforward somehow*" (P17), "*I have briefly tried meditation apps... I just kept not focusing. So [it] didn't really go that well.*" (P8)), a **lack of enjoyment** ("*I downloaded a few apps, but again, I find it annoying.*" (P17)), **accessibility issues** such as cost ("*I used Headspace [a meditation app], but I don't really do it anymore because I didn't want to pay the subscription*" (P14)), and **start-up work** ("*it's a lot of work that you need to get in touch with [a] therapist, and you need to experiment with them to see if you work well together.*" (P19)). Others had no particular strategies at all to dealing with stress in their everyday, such as P19 who said they "*just wait till [the stress] is gone*".

4.1.2 Perceived ineffectiveness of current stress reduction habits. When describing their current experiences of everyday stress, participants generally did not feel that they had good techniques for stress relief (e.g. "*I don't think I've had a ton of success*" (P3)). In particular, **participants felt they lacked self-sufficient skills**. For example, P16 said that their tendency to use distraction to deal with stress was "*evasive and destructive, which has worked enough, but I feel like it relies on something else... And I do think I don't have the techniques or knowledge to just by myself meditate or calm down*" (P16). When speaking about how she copes with making calls to customer service, a source of stress for her, P19 expressed, "*If I need to make the call by myself, I think I don't have a really good way to do that*" and this caused even more stress around not having a good way to deal with that stress: "*I think I feel stressed by not having a good way to cope with [the stress]*" (P19).

Additionally, heightened emotions was cited as a barrier for handling acutely distressing or stressful situations, such as interpersonal conflict. For these participants, calming their emotions was often seen as a necessity for them to them properly destress or self-soothe: "*I have to calm down a little bit. My reaction is a lot of anger. I clench my teeth, then I start sweating [and] my eyes become a little hazy...I do nothing until I've calmed down*" (P6). Another participant expressed, "*Anger really makes [it] hard to get out of [stressful] situations... it might be good for me to practice how to control my feelings by practicing breathing"* (P11).

4.2 Overall Perspectives on Using Simulation for Self-Care

After understanding participants' current experiences of stress, we proceeded to gather their thoughts on our prototypes and overall perspectives on the role that social simulation may take in their self-care.

Overall, participants responded very positively to some or all of the prototypes shown as well as the general idea of using virtual simulation to practice stress relief. **17 out of 19 participants said they would use one or more of the shown prototypes** (VR, AR, or text-based) in their current lives if they had access to the technology. Participants felt using simulation would **improve the outcomes** of situations ("*it would easily complement what I would normally do in reality, and just make all these situations work out better for me*" (P9)), help **explore different responses** ("*Just being able to articulate it beforehand and think, "okay, maybe I should try saying this instead". I think it would be helpful for me* (P18)), make them **feel more prepared** ("*actually having this scenario completely play out before it plays out* (P5)") and **mitigate negative emotions** ("*you can use it to mitigate some anger, stress*" (P9)). People even re-imagined the use

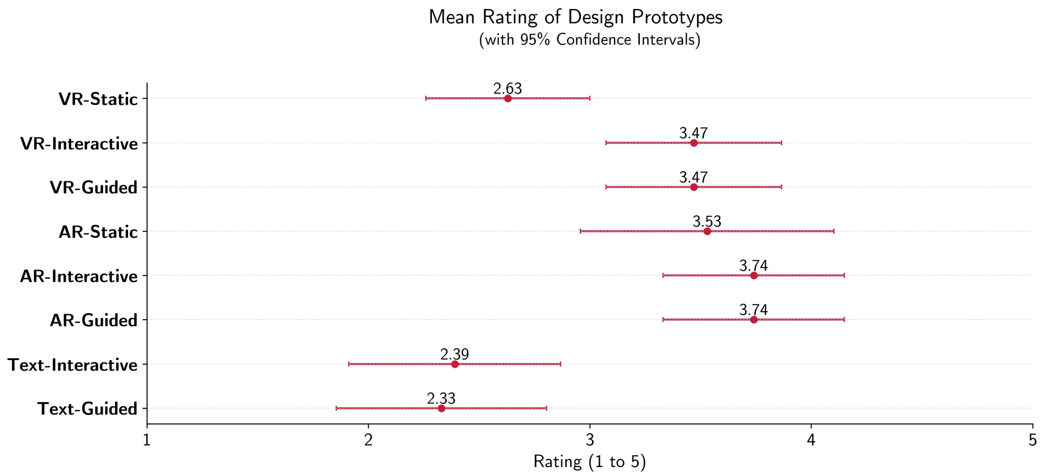


Fig. 6. Mean rating among all participants for our 8 design prototypes, with 95% confidence interval shown.

of this technology, such as P16 who mentioned that they feel it would be useful in professional therapeutic settings rather than just for personal self-care, saying, "*If a therapist had this available during a session, it could be quite useful to test-drive situations or problems that you're addressing during therapy.*".

4.2.1 Benefit of realistic self-care practice. Participants felt that the main benefit of social simulation was to practice stress relief techniques they could use "in-the-moment" of real-world scenarios. Doing so within a realistic environmental setting was praised by participants and seen as a step up from their current ways of coping with stress, as P16 notes, "*The closest that I'd imagine doing before I came into the study would have been practicing something in my head or in front of the mirror, which compared to this, is way less evocative and doesn't put you into the same mentality. So it wouldn't be as useful or as effective as something like this*". P5 also appreciated the practice "*of tackling the stress by just being in the situation itself*" as it helped them "*mitigate the fear*" when approaching an interpersonal conflict. In contrast to retroactively reflecting, the simulation helped surface emotions in-the-moment of stress, which could result in a more accurate and reflective picture of one's emotional state. For example, one participant noted that "*I feel like this one, I can do it in the middle of whatever I'm experiencing. I definitely feel more in control because other techniques, you just reflect on it afterwards. With this one, I feel more like reflecting in that moment, trying to observe my reaction. In that sense, it was more helpful*" (P14). Moreover, participants noted that doing the breathing guidance served as a reminder to calm themselves "*in the middle of a situation*" when they expected to be too emotionally elevated and "*may not really come up with that idea [to] breathe at this moment*" (P11). Our simulation was also seen as fostering better transferability of these skills to real-life situations compared to interventions that remove individuals from the context of stress. Participants enjoyed the practical aspect of practicing deep breathing while "in" the stress-inducing environment, versus other technological interventions for self-care that either have no environmental immersion [50] or use abstract or "escapist" (e.g. forest, beach) visuals [64, 78]:

> "*There's two types of things, right? There's one thing that you do on a daily basis just so that you don't get triggered as much...I feel like meditation falls in one of those, whereas deep breathing or stress balls or whatever people do – those fall into in-the-moment stress. And I feel like this is great for the second part...I*

> *could see myself using this, where I know a situation is probably going to trigger some stress for me. I would like to take all of those layers out as much as I can [by] using something like this before I actually enter into that situation.*" (P12)

4.2.2 Building Confidence and Preparedness. Participants also found the non-guidance versions (e.g. **`VR-Interactive`**, **`AR-Interactive`**, **`Text-Interactive`**) useful for simply preparedness and confidence having already had (virtual) experience with a situation beforehand. P18 articulated that this practice "*has the use of getting over the initial nervousness of confrontation*" and helped "*[get] over anxiety and trepidation*". P15 mentioned feeling more prepared after just using VR-Interactive without any mental health guidance: "*I would feel more prepared to actually confront them. The [avatar] showed me some attitude and I was kind of caught off-guard. If my roommate actually did that, I would have been prepared*". The technology also helped de-sensitize to anxious situations as one participant notes regarding her experience coping with social anxiety through avoidance: "*you start avoiding places that make you uncomfortable...but some stress you can't avoid*" and described using the simulation as a "*choose-your-own-adventure*" tool to "*practice stepping out of your comfort zone*" and to help her "*reintroduce to social situations*" (P13).

4.3 Modality: Benefits and Potential Risks of Immersion and Realism

We summarize below how participants responded to different modalities and allowances for immersion in our prototypes. Overall, participants expressed a desire for increased realism in the prototypes but also raised concerns about emotional safety.

4.3.1 Preference for Augmented Reality. In terms of which modalities were preferred among the eight prototypes, participants generally preferred AR over VR or text-based prototypes, with **`AR-Interactive`** and **`AR-Guided`** receiving the highest average ratings (3.74 out of 5) (see Figure 6). However, rather than a strict favorite, some saw value in multiple designs depending on context:

> "*If I'm out in a public scenario, and putting on the VR headset is a little much...then I started typing away. I think that's the best use for [the text-based chatbot]. For me, the AR one [if] I'm in the area in which I'm going to have the talk... The VR one [if] I'm inside, I'm alone.*" (P9)

> "*Using the VR is like the initial part, just to only be able to focus on the conversation. The AR is the next step where you're in the space where the issue is...so [first] being able to practice in a controlled environment, which is the VR, then move to the AR.*" (P18)

One notable consensus among participants regardless of their overall preferences was that **AR was most useful when available in the actual locations where they anticipated stressors**, such as the classroom they expected to give a talk in or their apartment if expecting a confrontation to occur in their home. For example, participants said that they'd "*use AR when it is close to where I will actually give a speech*" (P19) and practice stress-relief for a roommate conflict "*in my actual room...maybe I can breathe and walk around while I breathe*" (P9). Indeed, "*a lot of the confidence comes from just being comfortable in your space*" (P12) and practicing situations in the expected environment was already what some participants said they would do now if preparing for something like public speaking. This would also allow for people to practice their own **physicality**. For example, P9 felt AR would "*encourage me to move around a little bit, do hand gestures*" and, more than the dialogue, P17 saw the simulation as more useful for his "*physicality, you know moving around...engaging with people, which approaches I would do. I could train that*".

Experiencing AR also opened up suggestions of different mental health guidance options according to people's personal needs, such as navigating their surroundings to reduce stress. For example, P11 felt he would use AR to practice in his real home being able to retreat to a different room to do deep breathing if overwhelmed at a social party. Similarly, people could also practice doing guidance rooted in visual stimuli in the environment as P12 mentions: "*I know a few people who do struggle with social anxiety or confrontation. I remember somebody had mentioned to me that...there's a point in their room that they look at and that helps them just go into zone of 'I need to breathe three times'. [AR] could translate to that...you can just look at the same spot, and it should prompt the subconscious breathing*".

4.3.2 Environmental Realism. Although past work in social simulation has focused on agent interactions to simulate a "real-world" setting [70, 77], we found that **environmental factors are critical in participants' experience of social simulation**; immersive simulation goes beyond just accurate interactions or behaviors. In particular, ambient sound and visual immersion were noticed and desired by participants. For example, multiple participants felt that the background sound in our prototypes was key for their feelings of presence ("*listening to the sound [of the party] made me anxious*" (P11), "*the sound was very important [for immersion]*" (P17)) and wanted even further sound immersion, such as emphasizing particular sounds ("*In real life, when you're stressed or anxious, certain sounds are emphasized, and it adds to the feeling of stress*" (P17)).

We saw the importance of environmental factors for realism when comparing users' reactions between AR/VR to the text-based simulations, which participants described as "*artificial*" (P11) and "*more like a tech simulation*" (P10) despite GPT prompting and dialogue style being consistent across VR, AR, and text-based interactions. P16 pointed out the stark difference between the visual experience of AR/VR versus text: "*I'm comparing it now to what I was experiencing during the AR/VR experience...it feels a little more clinical, removed. Obviously it's the same engine and it's the same conversation model, but the act of either being in AR or VR space and having, albeit rudimentary, a representation of another person does help the believability and the immersion*". The disconnect participants felt in text-based simulations despite the same dialogue prompting as in VR/AR showed the importance of creating rich multisensory environments.

Participants consistently expressed a need for customizable avatars and environments to better reflect real-life scenarios: "*This scenario could be more personalized. For instance, if I could say, 'I'm attending this party tonight and meeting these specific people, and I'm nervous about it,' then maybe the conversations could reflect that*" (P14). Participants desired for simulations that not only felt plausible but were specifically relevant to their own lives. Even smaller physical differences, like positioning of avatars, could break immersion. For example, P12 pointed out that the roommate conflict simulation felt unnatural because the avatar stood facing across the user. She commented that she would not imagine this positioning for a conflict situation and desired the ability to customize this, stating, "*'do you want to be standing up or sitting down?' – maybe just as simple as that. The room that you have right now, I would choose to sit on the couch*". Generally, we found that participants desired modification features for avatars and environmental details to match their real-world expectations.

4.3.3 Risks in Realistic Simulation of Stress. While participants found our prototypes to generally be immersive, **they wanted the simulation to produce higher stress levels** to more closely emulate real life. Participants saw this as giving "*insights [on] how to handle or how I would feel...Basically, it's a way to anticipate my stress or my anxiety in a real setting*" (P17) and avoiding getting false security since "*[if you] go to practice with this, and then you go for a real environment with real people – maybe they [will not] be so gentle with you*" (P7). P10 echoed this and surfaced a potential limitation with using GPT for interactive dialogue, saying, "*I went in with the impression that no matter what I responded with, there would be people bringing up other new topics and things that I could respond to. In real-life conversation you're* *also obligated to keep the conversation flowing*". In fact, many participants correctly identified our dialogue engine as GPT-generated, and commented on GPT being unrealistically polite. Indeed, several participants chose not to use the available guidance during **`VR-Guided`**, citing a lack of stress. Although some felt guidance would be useful, they did not need it during the simulation: "*If the roommate would have persisted a little more, I would have probably [selected the breathing guidance]...maybe I just wasn't stressed enough in this situation*" (P12).

However, the risk of simulating high levels of stress is **potential risk of trauma**. For example, when P19 was asked whether she would like the simulation to be more realistic, she expressed hesitation and stated, "*I sometimes feel like I would be traumatized when I'm having conflict with somebody. I just want it to happen once. And I think practice will help to reduce the stress, but I kind of feel like maybe I'll get traumatized*". Another participant noted, "*If somebody has a very deep fear of public speaking, this might be too realistic...confronting the fear head on, depending on the person, might help or might not*" (P5). Some participants felt lacking some realism created a safer environment for practice. P8 expressed that greater realism "*kind of scares me*" and P5 appreciated the current balance as helpful but safe: "*I do like that it's not completely realistic. It helps me remember that it's just a simulation and that it's more for mental health*". The current level of (or lack of) realism allowed users to feel safer in verbal expression and more willing to engage in difficult scenarios. This allowed participants to feel it was "*okay to fail in the conversation*" (P14) and "*it's a safe environment... I know it's virtual [so] I am more relaxed...it will not have a repercussion*" (P17).

4.3.4 Value in Low-Immersion Designs. Despite a preference for immersive features, participants found value in low-immersion, text-based prototypes due to their accessibility and ease of use. The ease with which users could access text-based simulations made them a practical alternative for stress-relief practice in everyday settings. Participants mentioned that accessing chatbots "*might be more accessible for most users*" (P15), and "*in your everyday, I think this is way easier to pull up on your computer or your phone...like VR is something that most people either don't have or don't have on them always*" (P8). As a result, while environmental immersion was an important factor to people finding simulation helpful for practicing stress relief, participants still valued non-immersive but accessible systems.

4.4 Interactivity

Interactivity through dialogue was preferred across the board among participants. When asked, all participants explicitly said they preferred **`VR-Interactive`** and **`AR-Interactive`** over the static counterparts due to the dialogue. However, despite people preferring interactivity, non-interactive designs were still seen as more valuable than current means of practicing stress relief.

4.4.1 Interactive Dialogue. Having interactive, realistic dialogue was key for participants in evoking accurate reactions and emotions. The interactive nature evoked genuine emotional responses, which participants felt surprised but pleased by, that fostered deeper engagement and self-realizations. For example, P11 said "*I didn't really think the headset [could] make me very stressful...I was very surprised*" and the dialogue even evoked hesitancy akin to how participants would feel in real life:

> "*Even though it's a simulation, you could hear [my] stuttering and the hesitation. I was trying to still not go in with too much anger or irritation, but then also trying to be as direct with the problem.*" (P18)
>
> "*Even now, I was stumbling over my words and I wasn't really sure what I would say. So this is a perfect scenario, where I could think of how to approach the situation delicately.*" (P5)

Participants commented on enjoying how they were prompted to give realistic responses. P12, who engaged in an interpersonal conflict situation, stated that they enjoyed how "*it was not like they suggest the solution upfront because I feel like most of the times, it doesn't go that same way*" and P10 felt questions received during the public speaking scenario were "*in-depth*" and "*elicit a very thoughtful response*".

Beyond just realistic dialogue, participants emphasized that they desired more nonverbal cues such as facial expressions and body language due to its relevance in real life. P14 pointed out that "*nonverbal behavior is what's causing a lot of the anxieties as well*" and P19 also felt so regarding public speaking, stating, "*I would prefer [the avatars to] have more facial expression when I answer some questions...some [audience members] will make some faces that make me really nervous...it also adds more stress to myself, but I would say it will make the scenario more realistic*".

All 19 participants favored dictation using our speech-to-text feature over typing on the VR keyboard due to ease of use and its similarity to real life interactions. However, since our designs did not implement text-to-speech, nearly all participants expressed wanting audible voices from avatars. As P19 notes, avatars speaking out loud rather than our design of showing dialogue in avatars' text boxes would help give "*a shorter time to process the information*" to create "*more realistic interactions*" akin to real life.

4.4.2 Movement and Interaction with Environment. Movement and interaction with environment were interactions that participants expressed wanting more of. As we reviewed previously, people enjoyed the thought of AR due to their closeness to the real environment they expected to encounter stress. Some of this desire came from expecting physical movement to play a significant role when feeling discomfort. People therefore felt that the fact they were stationary in our VR/AR prototypes was limiting. For example, P13 noted that their natural response to feeling socially anxious would be to move around, "*my impulse would be to walk to a new location within the space. A lot of people in a social situation, if they see a snack table will be like, "look, snacks!" and you run over there.*" The ability for greater levels of interaction with the environment could help users also simulate different ways of handling stress, such as finding more comfortable spots in a room or occupying themselves with other objects. Additionally, allowing interaction with the avatars based on the simulation's environmental factors could help people practice icebreakers and conversational skills in a more natural way, as P13 points out about how socially anxious people may choose to converse when feeling nervous in social settings: "*it's the cat, or these plants are nice. You look for some shiny objects and start a conversation based on artwork in the room or something*".

4.4.3 Value in Non-Interactive Simulation. Users enjoyed the realism and interactive aspects of the prototypes, as reviewed above. However, interestingly we found that participants still perceived value even in the prototypes without any interaction (**`VR-Static`** and **`AR-Static`**) as they were still perceived as more helpful than participants' current means. For some, this was due to the fact that they added any kind of visual to an already existing simulation people currently do in their heads to prepare for difficult situations ("mirror the act of pulling it up in your head" (P16)). Similarly, regarding **`VR-Static`**, P12 felt, "*I'm comparing it to practicing in the mirror...it's the same fundamentals. You're trying to get used to seeing things that you're not normally used to saying...it is good*". The lack of realistic graphics was also not seen as necessary to still gain value from the technology for most participants: "*I don't think it's a problem that it's low poly or not necessarily realistic graphics. The act of being in this space and having a roommate character in front of you makes it into the psychology of being in that situation*" (P16).

4.5 Guidance Needs

Although participants still found the prototypes without mental health guidance useful, breathing guidance in VR-Guided and AR-Guided prototypes yielded an opportunity for **emotional regulation** and **more productive responses**.

Breathwork helped participants calm themselves and self-reflect on their emotional state. For instance, participants noted breathwork "*definitely made me calm down*" (P14) and helped "*mitigate some anger, stress*" (P9). Participants usually triggered breathwork when dialogue became more uncomfortable, such as a roommate becoming defensive or receiving a pressing question from an audience member. In doing so, participants were able to be "*more aware of the situation and gives me some time to actually watch what kind of emotions that I have in myself*" (P14) and that breathing during a more stressful encounter "*actually did a good job in helping manage mostly anger*" (P9).

Additionally, many participants also mentioned its usefulness for helping them respond more productively. For instance, participants stated "*with that breathing exercise, it will help me to be calm and think about the answer*" (P17) and "*being able to articulate it beforehand with a response and think, "okay, maybe I should try saying this instead*" (P18). Pausing during stressful moments gave a moment to "*recenter and get back to what I practiced*" (P18) as a healthier response to unexpected stress arising. Additionally, given participants' current methods of stress relief revolving around avoidance and distraction (see Section 4.1), it was particularly notable that participants felt that practicing breathing would help them to not immediately leave situations in unproductive ways. For example, P9 commented "*if I didn't press the breathing bubble, I would have just been like, 'okay, it's whatever' and then left when in reality, that's not going to resolve a situation. I think taking that extra time to breathe there was better*".

4.5.1 Transferable Skills to the Real-World. While some participants felt the practice was good for habit-building for applying to real-life (e.g. "*you can make it a practice...do it subconsciously so that it calms you down*" (P12)), other participants expressed that deep breathing was not necessarily a transferable practice due to its **impracticality during real-world situations**. As a result, we discovered that **transferability, practicality, and even social acceptance** were key dimensions that affect whether mental health guidance is perceived as useful. For instance, participants expressed hesitation on doing box breathing in real life:

> "*I do kind of worry that in real life people are not gonna wait for me to breathe and breathe out.*" (P14)
>
> "*People will think: why the long pause?*" (P6)
>
> "*I can't stop this conversation anytime without worrying about [the avatar's] feeling. So I can't just roughly [and] abruptly stop talking with him without answering his question.*" (P11)

As a result, people wanted and suggested different types of guidance that they thought were more relevant to their specific lives. For example, P19 suggested that her real-life calming mechanism during public speaking was seeing encouraging facial expressions from the audience and thus recommended simulating this. P18 mentioned techniques that felt easier to do in the middle of a conversation, such as, "*closing your eyes and counting to a certain number and then gather your thoughts and recenter yourself*" (P18). Other participants simply just wanted options, such as "*not only the breathing, but maybe like a couple other options I could click for how to relieve the stress*" (P9), or what they felt would be more relevant to the context: "*The assistant was being sarcastic and the guidance said to practice box breathing, but I don't think I needed box breathing for that. Maybe if they were being a little bit more aggressive, I would need box breathing*" (P15).

4.5.2 Manual Guidance: Empowering users' agency and self-reliance. We also investigated participants' preferences between GPT-timed breathing guidance and self-triggered guidance. One reason people preferred manual control was their confidence in recognizing their own stress. For example, one participant commented on how she understood her needs when the moment arose: "*It was kind of vague in the beginning whether I should turn [breathwork] on or not. But when the moment came that I felt, "oh, maybe this is a good time to take a break"...I just made more [of a] conscious decision on when I need[ed] that*" (P14). Participants also felt manual control better reflected real-life decision-making, encouraged habit-building, and facilitated learning. As P9 explained, "*It's more realistic and applicable if I control it, because in real life I would be controlling when I do something like that*" and P16 echoed, "*Realistically, you would be the one trying to execute your own calming techniques...I think you yourself are the best arbitrator of when something's stressful.*"

More explicit concerns over **intrusiveness** also arose. P9 commented, "*I wouldn't like it if...the simulation made me do the breathing thing, because maybe I didn't feel like I needed it. Or I felt like I could handle the situation without doing the breathing*" while P7 similarly felt "*forcing someone in middle of the conversation might make things much worse...what if I don't want to take a break?*". These automated recommendations could even cause people to feel self-conscious; P14 encountered a suggestion for breathing guidance during **`Text-Guided`** and stated, "*I felt like I wasn't feeling stressed, but I think I got direction that I'm stressed...Like, am I stressed? When I'm not stressed, do I look stressed? I don't really like being told that I look stressed*".

Interestingly, participants also expressed a general distrust in the accuracy and universality of AI-driven mental health guidance. Even participants who found GPT's recommendations accurate during the study remained skeptical about future scenarios. For example, P9, who found the breathing suggestion useful, still had reservations: "*I actually thought the system did pretty good on this one, I thought the situation was getting a little stressful...but I don't know if it will do that every time. But for this time, it did good*".

4.5.3 Automated Guidance: Enhancing Self-Awareness. On the other hand, automated guidance helped users reflect on their emotional state and thus could lead to gaining self-awareness. While some participants expressed concerns about system-determined guidance, others appreciated its value in moments when stress might go unnoticed. P19 shared that she preferred LLM-powered guidance suggestions due to a lack of anticipation of her own stress, stating, "*When I was answering some of the questions previously, I did not feel like I'm stressed until I was stuck. Then I started to feel stressed, and that's something that I did not anticipate*". For these participants, automated guidance also was more intuitive to include in a training system such as ours in order to develop further self-awareness skills and allow users to consciously notice moments of stress they may have otherwise missed.

Overall, we found that users feel automated guidance can play a supportive role in helping users develop better self-awareness and emotional regulation skills. However, serious concerns over accuracy and agency also emerged. Perhaps due to the distrust of LLM-powered recommendations and its universality, multiple participants surprisingly suggested using **biosignals** to guide automatic recommendations instead. In subsequent interviews, when probed with the idea, all participants reacted positively to this idea. For example, P7 commented that "I prefer to take charge. But I guess for such a system, tracking your heart or something for it to be accurate? Or maybe your voice tone." and P16, another participant who suggested this idea, said, "I like the idea of monitoring your vitals and signs of stress, maybe especially in a VR/AR setting. Having like a non-intrusive pop-up or gentle reminder" when the user showed a signal like high heart rate. Integrating biosignal monitoring could also provide an alternative route to an accurate and personalized approach that still facilitates practicing self-awareness around stress management.

5 DISCUSSION

This study explored user perceptions of applying social simulation technologies for practicing self-care techniques for their everyday situations. Through prototype-driven interviews with 19 participants using eight built prototypes, we gathered valuable insights into the benefits, risks and challenges, and future design implications for applying these technologies to self-care.

5.1 Designing Simulation for Everyday Self-Care

As Nunes et al. note in a review of HCI technology for self-care, the concept of technological support for people's mental health should be expanded to center on people's everyday life experiences rather than just patients of chronic conditions [65]. Our study contributes to the field of HCI technology for everyday self-care and explored what design needs exist in this relatively unexplored space.

Modality. One of the most significant findings was the clear preference for AR among our prototypes. Participants envisioned using AR in real-world environments to overlay both practice scenarios and mental health guidance, allowing for the direct application of coping strategies in the settings they encounter in daily life. In the context of everyday stress, where individuals are frequently exposed to environments that trigger emotional or mental distress, our interview findings suggest that AR lowers barriers in transferring skills to real-world settings and individuals prefer to practice coping techniques in spaces that closely resemble real-world settings. While high levels of immersion in VR can foster engagement [] and be beneficial for simulating scenarios that aren't immediately accessible, its separation from real-world contexts may limit the practical application of skills learned in the virtual space. Text-based prototypes, unsurprisingly, were considered the least immersive given absence of visual and auditory feedback. However, many participants said they would use text-based systems because of its availability throughout the day, suggesting that even non-immersive tools have a valuable role in the self-care space.

Interaction. Participants clearly favored features like dialogue and environmental interaction, as these allowed them to practice communication skills in a system-responsive manner. This interactivity provided more opportunities to experience realistic stress and practice self-soothing techniques in response. Notably, even non-interactive prototypes like VR-Static and AR-Static were considered more helpful than participants' current methods of stress relief, indicating that while interactivity is desired, it is not essential for perceived usefulness.

Participants' reactions to our prototypes echoed past findings [] that LLMs are capable of generating natural, context-aware dialogue that participants found realistic compared to real-world interactions. In the future, given the success of speech-to-text in our study, LLMs could be used to also dynamically design environments rather than just provide dialogue as it did in our study. Given the accuracy of our speech-to-text engine using WhisperAI, it is plausible that users could verbally describe anticipated situations, and the LLM could then generate corresponding dialogue, character personas, physical avatar appearances, and environmental features to create a VR or AR scene for the user to practice with. Since participants consistently expressed a desire for customization features to adjust simulations to reflect their personal environments and stressors, using LLMs for designing more aspects of the simulation could facilitate this capability while reducing the time, effort, and technical skill needed on the user's part.

Mental Health Guidance. As seen through our interviews in how people responded to box breathing and their original approaches to self-care in the initial parts of our interview, there is not a one-size-fits-all technique for in-the-moment stress relief. Overall, designs should opt for flexibility or adaptability in guidance. In terms of our specific system, though, participants found the guided interventions useful during moments of heightened stress. However, an important theme that emerged was the transferability of such techniques to real-world interactions. Participants expressed concerns about the practicality of taking a moment to themselves during a heated conversation or stressful social event. Although one method of addressing these concerns is to offer techniques that are more easily transferable, this feedback might also suggest a greater opportunity to also teach people how to advocate for their own self-care needs in social settings and/or learn to transition into self-soothing skills during social settings. For example, instead of social simulation pausing the scene when the breathing guidance was triggered as in our study's designs, social simulation could also continue the interactions while guidance appeared and even simulate realistic ways that people may respond during these moments (e.g. facial expressions, asking a question again if the user does not respond in a timely way). By doing so, users can practice interaction necessary to pause, transition into self-soothing techniques, or step away from a conversation to manage stress. While our prototypes focused on the concept of individuals learning these skills to incorporate into their own daily lives, our interviews also surfaced the significant role that social dynamics play in the application of self-care strategies and even the interpersonal consequences of doing so.

5.2 Ensuring Safety and Agency through Customization

Although our simulation is designed as a safe environment to help users practice difficult situations, simulation can also carry risk in itself. The most apparent risk that emerged through our interviews was a tension between people wanting more realism, but also having concern over the potential of discomfort and trauma. Balancing realism and virtual imagery is crucial, as it impacts both user comfort and the risk of trauma. As one participant had noted in our study, stressful situations like interpersonal conflict can be difficult enough to deal with already just once in the real-world.

Our recommendation for giving users more control, minimizing surprises, and preventing stress responses beyond users' boundaries is to implement several **customization features** in the simulation. Customization features were mainly suggested by participants as contributing to realism, but that same control can ensure safety for users' boundaries as well. Customization of avatar and environmental features, for example, can be made more real or artificial according to users' preferences. Since sound was a key immersive element as well, simple features like volume control or the option to mute should be available at anytime during simulation in order to reduce auditory immersion if needed. Additionally, a theme that emerged from our findings was *tailoring of difficulty level*, given that some participants wanted more challenging situations to practice with while others had concerns over trauma. Adjustments in mental healthcare to gradually lead up on more challenging situations is well-established in traditional therapeutic settings already, such as "gradual exposure" in Cognitive Behavioral Therapy [20] and traditional Exposure Therapy [21]. Challenge in dialogue, such as a roommate being more defensive or combative in our simulation, may be tailored to GPT, and further work may need to be done to find prompting techniques that generate appropriately easy versus difficult output that is appropriate and consistent for users. As a result, we strongly recommend simulation for self-care provide control (and thus more safety) over specific features that may affect one's stress levels (e.g. amount of people at a party, volume levels) as well as difficulty when prompting LLM-powered dialogue. Overall, we find customization not only gives users agency, but also mitigate risks as users navigate their own boundaries.

As one participant suggested in our study, social simulation has potential to go beyond personal care and could be deployed in a clinical or therapeutic setting for supervised observance. Doing so may help therapists or clinicians see first-hand how their clients react in situations they encounter in life outside therapy, suggest coping techniques, and allow clients to reflect more accurately on feelings during situations they bring up in therapy. Although our work was originally built with a user's own individual, non-supervised learning and practice in mind, it is plausible that learning unsupervised may be best for people who already have personal understanding of their stress [65]. Different autonomy levels may be required for different "treatments" – while some care decisions such as those people make in our study's scenarios (e.g. public speaking, conflict) are done on a somewhat regular basis, other situations causing higher risk distress decisions should be supported and supervised by professionals.

We also note that participants had very positive reactions to incorporating biofeedback into the simulation system. Using biofeedback through heart rate monitoring or skin conductance measurements [37] could inform the system when the user is becoming stressed (and even adjust the simulation accordingly) but also generally make the user more informed of their physiological state for them to gain more self-awareness. Past work has often applied HCI technology, including wearable devices and biofeedback, just for informative purposes for users' health as symptoms and trends can often go unnoticed [65] and have become popular inclusions for mental health technology in recent years to help guide people to understand their habits and responses as well as the characteristics of their environment that contribute to their status. Given people's concerns over LLM-powered recommendations' accuracy but their desire for more self-awareness and reflection opportunities, future iterations of social simulation should consider using people's biofeedback to calibrate system responses and generally connect the user more with their own body's signals during mental health and self-care interventions.

5.3 Shifting Paradigms in HCI for Mental Health

5.3.1 Designing for Self-Care in the Real-World. Prior literature that has tackled immersion for mental and emotional well-being have placed the user in the beach [14, 106], gaming environments [90], abstract scenes [64], and outer space [62]. Escapism has so far been the traditional way that HCI research has approached helping people find emotional relief, transporting users to a different landscape as a temporary escape from their current mental or emotional difficulties. Escapist approaches certainly have their place in stress reduction – studies about escapism in gaming for mental health have found that these approaches can indeed help with emotional regulation if not used to an extreme [48]. However, there is also evidence to suggest that distraction, which was our participants' most common method of coping, can be effective as a short-term strategy for mood regulation but is not sustainable for long term well-being nor better for well-being than more active emotional regulation strategies [10]. Using escapism as coping can also cause more feelings of isolation, while methods to emotionally regulate directly (such as assessing emotions while in distress) can lead to better well-being [48]. Our findings too suggest a significant gap and need for reflecting the real-world and its stressors in order to create applicable, transferable, and sustainable skills that we can apply without technological involvement in our everyday lives. In this way, our work attempts to shift the conversation from **how to apply technology so it can make people feel better, to how to apply technology so it can *teach people how* to feel better**. Realism played a crucial role in participants positive experiences during our study, and participants consistently favored tools that replicated real-world environments (e.g. AR, desiring more realistic body language interactions). Immersion enabled users to apply coping mechanisms in settings that mirrored the actual environments where stress arises, such as their homes or workplaces; participants felt that this design ultimately brought more confidence to encountering stressful situations as well as better transfer of emotional regulation and reflection skills to real-life situations.

5.3.2 Reframing HCI for Self-Care. Lastly, we discuss how we hope our work contributes to the broader landscape of HCI for mental health by framing self-care — and mental healthcare in general — as **fundamental needs for everyone**, not just those with diagnosable, chronic conditions. Traditionally, self-care technologies in HCI have been positioned within the framework of clinical interventions or therapeutic support, often focusing on structured tools for chronic conditions like depression or anxiety []. However, as our findings highlight, there is a growing need to reframe self-care in HCI to create tools for everyday practice; in particular, we aimed to use HCI methods to equip users with the skills to manage stress in real-time for the contexts where it occurs. Our work's findings and implications also emphasize the idea of people's own **self-efficacy** in taking care of themselves. As seen in our interviews, people want to be self-sufficient but do not currently engage in means to do so when it comes to mental health, despite it being crucial for our well-being just as much (and maybe even more) as social or professional support needs. In our study, participants were able to engage in simulation that they felt gave agency to learning self-care skills in the long-term to enact on their own in the real-world.

In addition, self-care in HCI has traditionally focused on reactive strategies to help people manage stress after it occurs [] or self-management and tracking for specific conditions and diseases []. However, our study suggests that HCI research can go beyond these framings to instead foster **proactive self-care** by allowing users to practice personal coping skills before they enter stressful situations. This shifts self-care from a purely reactive measure to an ongoing skill-building process, where users can refine their emotional regulation techniques over time. HCI's role in supporting this proactive approach is critical, as it positions technology not just as a fix but as an integrated part of daily life that empowers users to actively manage their mental health on their own terms.

6 CONCLUSION

Our study highlights the potential of social simulation to support self-care practices. By exploring participants' responses to prototypes utilizing virtual reality (VR), augmented reality (AR), and large language models (LLMs), we examined how people react to and envision social simulation for practicing stress relief for their everyday lives. Our findings reveal that people largely lack self-sufficient skills to self-soothe during their everyday contexts, and practicing self-care methods during realistic simulation of these real-world environments encouraged habit-building and emotional regulation. We also uncover users' concerns and trade-offs in different design decisions such as threats to user agency with LLM-powered recommendations for mental health guidance and possible risks of trauma with hyper-realistic simulation. Ultimately, we provide a first look at how simulation can provide a safe, virtual environment for people to practice self-care, filling a critical gap in HCI for mental health.

REFERENCES

[1] Javier Alvarez-Galvez. 2016. Network models of minority opinion spreading: Using agent-based modeling to study possible scenarios of social contagion. *Social Science Computer Review* 34, 5 (2016), 567–581.

[2] American Psychological Association 2023. *Stress in America 2023: A nation recovering from collective trauma.* Retrieved Sept 1, 2024 from https://www.apa.org/news/press/releases/stress/2023/collective-trauma-recovery

[3] Nazanin Andalibi, Oliver L Haimson, Munmun De Choudhury, and Andrea Forte. 2016. Understanding social media disclosures of sexual abuse through the lenses of support seeking and anonymity. In *Proceedings of the 2016 CHI conference on human factors in computing systems*. 3906–3918.

[4] Petter Bae Bae Brandtzæg, Marita Skjuve, Kim Kristoffer Kristoffer Dysthe, and Asbjørn Følstad. 2021. When the social becomes non-human: young people's perception of social support in chatbots. In *Proceedings of the 2021 CHI conference on human factors in computing systems*. 1–13.

[5] Luke Balcombe and Diego De Leo. 2022. Human-computer interaction in digital mental health. In *Informatics*, Vol. 9. MDPI, 14.

[6] Jakob E Bardram, Mads Frost, Károly Szántó, Maria Faurholt-Jepsen, Maj Vinberg, and Lars Vedel Kessing. 2013. Designing mobile health technology for bipolar disorder: a field trial of the monarca system. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 2627–2636.

[7] Tanya GK Bentley, Gina D'Andrea-Penna, Marina Rakic, Nick Arce, Michelle LaFaille, Rachel Berman, Katie Cooley, and Preston Sprimont. 2023. Breathing practices for stress and anxiety reduction: Conceptual framework of implementation guidelines based on a systematic review of the published literature. *Brain Sciences* 13, 12 (2023), 1612.

[8] Cristina Botella, Javier Fernández-Álvarez, Verónica Guillén, Azucena García-Palacios, and Rosa Baños. 2017. Recent progress in virtual reality exposure therapy for phobias: a systematic review. *Current psychiatry reports* 19 (2017), 1–13.

[9] Virginia Braun and Victoria Clarke. 2012. Thematic analysis. (2012).

[10] Patricia C Broderick. 2005. Mindfulness and coping with dysphoric mood: Contrasts with rumination and distraction. *Cognitive therapy and research* 29 (2005), 501–510.

[11] Kevin J Bruggeman and Skylar W Wurster. 2018. The hiatus system: virtual healing spaces: low dose mindfulness based stress reduction virtual reality application. In *ACM siggraph 2018 appy hour*. 1–2.

[12] Eleanor R Burgess, Kathryn E Ringland, Jennifer Nicholas, Ashley A Knapp, Jordan Eschler, David C Mohr, and Madhu C Reddy. 2019. " I think people are powerful" The Sociality of Individuals Managing Depression. *Proceedings of the ACM on Human-computer Interaction* 3, CSCW (2019), 1–29.

[13] Jerome Carson and Elizabeth Kuipers. 1998. Stress management. *Occupational stress: Personal and professional approaches* (1998), 157.

[14] Amaya Chandrasiri, James Collett, Eric Fassbender, and Alexander De Foe. 2020. A virtual reality approach to mindfulness skills training. *Virtual Reality* 24, 1 (2020), 143–149.

[15] Szu-Wei Cheng, Chung-Wen Chang, Wan-Jung Chang, Hao-Wei Wang, Chih-Sung Liang, Taishiro Kishimoto, Jane Pei-Chen Chang, John S Kuo, and Kuan-Pin Su. 2023. The now and future of ChatGPT and GPT in psychiatry. *Psychiatry and clinical neurosciences* 77, 11 (2023), 592–596.

[16] Andrea Chester and Carolyn A Glass. 2006. Online counselling: A descriptive analysis of therapy services on the Internet. *British Journal of Guidance & Counselling* 34, 2 (2006), 145–160.

[17] Jaewoo Chung, Francesco Pagnini, and Ellen Langer. 2016. Mindful navigation for pedestrians: Improving engagement with augmented reality. *Technology in Society* 45 (2016), 29–33.

[18] David M. Clark, Paul M. Salkovskis, and A.J. Chalkley. 1985. Respiratory control as a treatment for panic attacks. *Journal of Behavior Therapy and Experimental Psychiatry* 16, 1 (1985), 23–30. https://doi.org/10.1016/0005-7916(85)90026-6

[19] Cleveland Clinic 2021. *How Box Breathing Can Help You Destress*. Retrieved Sept 9, 2024 from https://health.clevelandclinic.org/box-breathing-benefits

[20] Judith A Cohen, Anthony P Mannarino, Matthew Kliethermes, and Laura A Murray. 2012. Trauma-focused CBT for youth with complex trauma. *Child abuse & neglect* 36, 6 (2012), 528–541.

[21] Michelle G Craske, Michael Treanor, Christopher C Conway, Tomislav Zbozinek, and Bram Vervliet. 2014. Maximizing exposure therapy: An inhibitory learning approach. *Behaviour research and therapy* 58 (2014), 10–23.

[22] Davide Crivelli, Giulia Fronda, Irene Venturella, and Michela Balconi. 2019. Supporting mindfulness practices with brain-sensing devices. Cognitive and electrophysiological evidences. *Mindfulness* 10 (2019), 301–311.

[23] Sanobar Dar, Aniko Ekart, and Ulysses Bernardet. 2022. The virtual human breathing coach. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. IEEE, 434–436.

[24] Munmun De Choudhury, Sachin R Pendse, and Neha Kumar. 2023. Benefits and harms of large language models in digital mental health. *arXiv preprint arXiv:2311.14693* (2023).

[25] Mairi Therese Deighan, Amid Ayobi, and Aisling Ann O'Kane. 2023. Social virtual reality as a mental health tool: how people use VRChat to support social connectedness and wellbeing. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–13.

[26] Morton Deutsch. 1994. Constructive conflict resolution: Principles, training, and research. *Journal of social issues* 50, 1 (1994), 13–32.

[27] Steven P Dow, Alana Glassco, Jonathan Kass, Melissa Schwarz, Daniel L Schwartz, and Scott R Klemmer. 2010. Parallel prototyping leads to better design results, more divergence, and increased self-efficacy. *ACM Transactions on Computer-Human Interaction (TOCHI)* 17, 4 (2010), 1–24.

[28] Karen Kangas Dwyer and Marlina M Davidson. 2012. Is public speaking really more feared than death? *Communication Research Reports* 29, 2 (2012), 99–107.

[29] Caroline J Falconer, Aitor Rovira, John A King, Paul Gilbert, Angus Antley, Pasco Fearon, Neil Ralph, Mel Slater, and Chris R Brewin. 2016. Embodying self-compassion within virtual reality and its effects on patients with depression. *BJPsych open* 2, 1 (2016), 74–80.

[30] Rachel R Feinberg, Udaya Lakshmi, Matthew J Golino, and Rosa I Arriaga. 2022. Zenvr: Design evaluation of a virtual reality learning system for meditation. In *Proceedings of the 2022 CHI conference on human factors in computing systems*. 1–15.

[31] Daniel Freeman. 2008. Studying and treating schizophrenia using virtual reality: a new paradigm. *Schizophrenia bulletin* 34, 4 (2008), 605–610.

[32] Daniel Freeman, Sarah Reeve, Abi Robinson, Anke Ehlers, David Clark, Bernhard Spanlang, and Mel Slater. 2017. Virtual reality in the assessment, understanding, and treatment of mental health disorders. *Psychological medicine* 47, 14 (2017), 2393–2400.

[33] Guanghui Fu, Qing Zhao, Jianqiang Li, Dan Luo, Changwei Song, Wei Zhai, Shuo Liu, Fan Wang, Yan Wang, Lijuan Cheng, et al. 2023. Enhancing psychological counseling with large language model: A multifaceted decision-support system for non-professionals. *arXiv preprint arXiv:2308.15192* (2023).

[34] Nigel Gilbert and Klaus Troitzsch. 2005. *Simulation for the social scientist*. McGraw-Hill Education (UK).

[35] Florian Grieger, Holger Klapperich, and Marc Hassenzahl. 2021. Trash it, punch it, burn it–using virtual reality to support coping with negative thoughts. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–6.

[36] Sandra R Harris, Robert L Kemmerling, and Max M North. 2002. Brief virtual reality therapy for public speaking anxiety. *Cyberpsychology & behavior* 5, 6 (2002), 543–550.

[37] Blake Anthony Hickey, Taryn Chalmers, Phillip Newton, Chin-Teng Lin, David Sibbritt, Craig S McLachlan, Roderick Clifton-Bligh, John Morley, and Sara Lal. 2021. Smart devices and wearable technologies to detect and monitor mental health conditions and stress: A systematic review. *Sensors* 21, 10 (2021), 3461.

[38] James D Hollan, Edwin L Hutchins, and Louis Weitzman. 1984. STEAMER: An interactive inspectable simulation-based training system. *AI magazine* 5, 2 (1984), 15–15.

[39] Toshiro Horigome, Shunya Kurokawa, Kyosuke Sawada, Shun Kudo, Kiko Shiga, Masaru Mimura, and Taishiro Kishimoto. 2020. Virtual reality exposure therapy for social anxiety disorder: a systematic review and meta-analysis. *Psychological medicine* 50, 15 (2020), 2487–2497.

[40] Zihui Hu, Hanchao Hou, and Shiguang Ni. 2024. Grow with Your AI Buddy: Designing an LLMs-based Conversational Agent for the Measurement and Cultivation of Children's Mental Resilience. In *Proceedings of the 23rd Annual ACM Interaction Design and Children Conference*. 811–817.

[41] Regina Kaplan-Rakowski, Karen R Johnson, and Tomasz Wojdynski. 2021. The impact of virtual reality meditation on college students' exam performance. *Smart Learning Environments* 8 (2021), 1–15.

[42] Eirini Karyotaki, Heleen Riper, Jos Twisk, Adriaan Hoogendoorn, Annet Kleiboer, Adriana Mira, Andrew Mackinnon, Björn Meyer, Cristina Botella, Elizabeth Littlewood, et al. 2017. Efficacy of self-guided internet-based cognitive behavioral therapy in the treatment of depressive symptoms: a meta-analysis of individual participant data. *JAMA psychiatry* 74, 4 (2017), 351–359.

[43] Sylvia Deidre Kauer, Cheryl Mangan, and Lena Sanci. 2014. Do online mental health services improve help-seeking for young people? A systematic review. *Journal of medical Internet research* 16, 3 (2014), e3103.

[44] Mina J Kian, Mingyu Zong, Katrin Fischer, Abhyuday Singh, Anna-Maria Velentza, Pau Sang, Shriya Upadhyay, Anika Gupta, Misha A Faruki, Wallace Browning, et al. 2024. Can an LLM-powered socially assistive robot effectively and safely deliver cognitive behavioral therapy? A study with university students. *arXiv preprint arXiv:2402.17937* (2024).

[45] Taewan Kim, Seolyeong Bae, Hyun Ah Kim, Su-woo Lee, Hwajung Hong, Chanmo Yang, and Young-Ho Kim. 2024. MindfulDiary: Harnessing Large Language Model to Support Psychiatric Patients' Journaling. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–20.

[46] David Kingdon. 2009. Everybody gets stressed … it's just the way we react that differs. *Psychiatric Bulletin* 33, 12 (2009), 441–442. https://doi.org/10.1192/pb.bp.109.025007

[47] Nikki Kiyimba and Michelle O'Reilly. 2020. The clinical use of Subjective Units of Distress scales (SUDs) in child mental health assessments: a thematic evaluation. *Journal of Mental Health* 29, 4 (2020), 418–423.

[48] Mehmet Kosa and Ahmet Uysal. 2020. Four pillars of healthy escapism in games: Emotion regulation, mood management, coping, and recovery. *Game user experience and player-centered design* (2020), 63–76.

[49] Merel Krijn, Paul MG Emmelkamp, Ragnar P Olafsson, and Roeline Biemond. 2004. Virtual reality exposure therapy of anxiety disorders: A review. *Clinical psychology review* 24, 3 (2004), 259–281.

[50] Oskari Lahtinen, Jenni Aaltonen, Johanna Kaakinen, Lena Franklin, and Jukka Hyönä. 2023. The effects of app-based mindfulness practice on the well-being of university students and staff. *Current Psychology* 42, 6 (2023), 4412–4421.

[51] Tin Lai, Yukun Shi, Zicong Du, Jiajie Wu, Ken Fu, Yichao Dou, and Ziqi Wang. 2023. Supporting the Demand on Mental Health Services with AI-Based Conversational Large Language Models (LLMs). *BioMedInformatics* 4, 1 (2023), 8–33.

[52] Emily G Lattie, Rachel Kornfield, Kathryn E Ringland, Renwen Zhang, Nathan Winquist, and Madhu Reddy. 2020. Designing mental health technologies that support the social ecosystem of college students. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–15.

[53] Mei Hui Lim, Vahid Aryadoust, and Gianluca Esposito. 2023. A meta-analysis of the effect of virtual reality on reducing public speaking anxiety. *Current Psychology* 42, 15 (2023), 12912–12928.

[54] Yuhan Liu, Anna Fang, Glen Moriarty, Robert Kraut, and Haiyi Zhu. 2023. Agent-based Simulation for Online Mental Health Matching. *arXiv preprint arXiv:2303.11272* (2023).

[55] Siyuan Brandon Loh and Aravind Sesagiri Raamkumar. 2023. Harnessing Large Language Models' Empathetic Response Generation Capabilities for Online Mental Health Counselling Support. *arXiv preprint arXiv:2310.08017* (2023).

[56] Yanick Xavier Lukic, Alvaro Hernandez Reguera, Amanda Cotti, Elgar Fleisch, Tobias Kowatsch, et al. 2021. Physiological responses and user feedback on a gameful breathing training app: within-subject experiment. *JMIR serious Games* 9, 1 (2021), e22802.

[57] Zilin Ma, Yiyang Mei, and Zhaoyuan Su. 2023. Understanding the benefits and challenges of using large language model-based conversational agents for mental well-being support. In *AMIA Annual Symposium Proceedings*, Vol. 2023. American Medical Informatics Association, 1105.

[58] Mounica Maddela, Megan Ung, Jing Xu, Andrea Madotto, Heather Foran, and Y-Lan Boureau. 2023. Training models to generate, recognize, and reframe unhelpful thoughts. *arXiv preprint arXiv:2307.02768* (2023).

[59] AH Maslow. 1943. A theory of human motivation. *Psychological Review google schola* 2 (1943), 21–28.

[60] Rahel Maué and Rahel Flechtner. 2022. "Hopohopo": a virtual reality awareness application about social anxiety disorder. In *Proceedings of Mensch und Computer 2022*. 420–424.

[61] Katharina Meyerbröker and Paul MG Emmelkamp. 2010. Virtual reality exposure therapy in anxiety disorders: a systematic review of process-and-outcome studies. *Depression and anxiety* 27, 10 (2010), 933–944.

[62] Noah Miller, Ekaterina R Stepanova, John Desnoyers-Stewart, Ashu Adhikari, Alexandra Kitson, Patrick Pennefather, Denise Quesnel, Katharina Brauns, Anika Friedl-Werner, Alexander Stahn, et al. 2023. Awedyssey: design tensions in eliciting self-transcendent emotions in virtual reality to support mental well-being and connection. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference*. 189–211.

[63] David C Mohr, Kathryn Noth Tomasino, Emily G Lattie, Hannah L Palac, Mary J Kwasny, Kenneth Weingardt, Chris J Karr, Susan M Kaiser, Rebecca C Rossom, Leland R Bardsley, et al. 2017. IntelliCare: an eclectic, skills-based app suite for the treatment of depression and anxiety. *Journal of medical Internet research* 19, 1 (2017), e10.

[64] Sandra Olga Ka Man Ng and Carlos Sena Caires. 2023. Virtual Serenity: Exploring a Chakra Meditation Dome in Virtual Reality for Stress Relief in Macao. In *Proceedings of the 11th International Conference on Digital and Interactive Arts*. 1–4.

[65] Francisco Nunes, Nervo Verdezoto, Geraldine Fitzpatrick, Morten Kyng, Erik Grönvall, and Cristiano Storni. 2015. Self-care technologies in HCI: Trends, tensions, and opportunities. *ACM Transactions on Computer-Human Interaction (TOCHI)* 22, 6 (2015), 1–45.

[66] David Opriş, Sebastian Pintea, Azucena García-Palacios, Cristina Botella, Ştefan Szamosközi, and Daniel David. 2012. Virtual reality exposure therapy in anxiety disorders: a quantitative meta-analysis. *Depression and anxiety* 29, 2 (2012), 85–93.

[67] World Health Organization. 2022. *WHO guideline on self-care interventions for health and well-being, 2022 revision.* World Health Organization.

[68] Maryann E Owens and Deborah C Beidel. 2015. Can virtual reality effectively elicit distress associated with social anxiety disorder? *Journal of Psychopathology and Behavioral Assessment* 37 (2015), 296–305.

[69] 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*. 1–22.

[70] Joon Sung Park, Lindsay Popowski, Carrie Cai, Meredith Ringel Morris, Percy Liang, and Michael S Bernstein. 2022. Social simulacra: Creating populated prototypes for social computing systems. In *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology*. 1–18.

[71] Rakesh Patibanda, Florian'Floyd' Mueller, Matevz Leskovsek, and Jonathan Duckworth. 2017. Life tree: understanding the design of breathing exercise games. In *Proceedings of the annual symposium on computer-human interaction in play*. 19–31.

[72] Darius Plikynas, Aistis Raudys, and Sarunas Raudys. 2015. Agent-based modelling of excitation propagation in social media groups. *Journal of Experimental & Theoretical Artificial Intelligence* 27, 4 (2015), 373–388.

[73] Claudette Pretorius, Darragh McCashin, Naoise Kavanagh, and David Coyle. 2020. Searching for mental health: a mixed-methods study of young people's online help-seeking. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–13.

[74] Mirjana Prpa, Kıvanç Tatar, Jules Françoise, Bernhard Riecke, Thecla Schiphorst, and Philippe Pasquier. 2018. Attending to breath: exploring how the cues in a virtual environment guide the attention to breath and shape the quality of experience to support mindfulness. In *Proceedings of the 2018 designing interactive systems conference*. 71–84.

[75] Greg M Reger, Patricia Koenen-Woods, Kimberlee Zetocha, Derek J Smolenski, Kevin M Holloway, Barbara O Rothbaum, JoAnn Difede, Albert A Rizzo, Amanda Edwards-Stewart, Nancy A Skopp, et al. 2016. Randomized controlled trial of prolonged exposure using imaginal exposure vs. virtual reality exposure in active duty soldiers with deployment-related posttraumatic stress disorder (PTSD). *Journal of consulting and clinical psychology* 84, 11 (2016), 946.

[76] Yuqing Ren and Robert E Kraut. 2010. Agent-based modeling to inform online community theory and design: Impact of discussion moderation on member commitment and contribution. *Second round revise and resubmit at Information Systems Research* 21, 3 (2010).

[77] Yuqing Ren and Robert E Kraut. 2014. Agent-based modeling to inform online community design: Impact of topical breadth, message volume, and discussion moderation on member commitment and contribution. *Human–Computer Interaction* 29, 4 (2014), 351–389.

[78] Joan Sol Roo, Renaud Gervais, Jeremy Frey, and Martin Hachet. 2017. Inner garden: Connecting inner states to a mixed reality sandbox for mindfulness. In *Proceedings of the 2017 CHI conference on human factors in computing systems*. 1459–1470.

[79] Marie-Laure Ryan. 1994. Immersion vs. interactivity: Virtual reality and literary theory. *Postmodern culture* 5, 1 (1994).

[80] Jessica Schroeder, Chelsey Wilkes, Kael Rowan, Arturo Toledo, Ann Paradiso, Mary Czerwinski, Gloria Mark, and Marsha M Linehan. 2018. Pocket skills: A conversational mobile web app to support dialectical behavioral therapy. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–15.

[81] Frank Schweitzer and David Garcia. 2010. An agent-based model of collective emotions in online communities. *The European Physical Journal B* 77 (2010), 533–545.

[82] Elizabeth Seabrook, Ryan Kelly, Fiona Foley, Stephen Theiler, Neil Thomas, Greg Wadley, and Maja Nedeljkovic. 2020. Understanding how virtual reality can support mindfulness practice: mixed methods study. *Journal of medical Internet research* 22, 3 (2020), e16106.

[83] Omar Shaikh, Valentino Emil Chai, Michele Gelfand, Diyi Yang, and Michael S Bernstein. 2024. Rehearsal: Simulating conflict to teach conflict resolution. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–20.

[84] Ashish Sharma, Inna W Lin, Adam S Miner, David C Atkins, and Tim Althoff. 2023. Human–AI collaboration enables more empathic conversations in text-based peer-to-peer mental health support. *Nature Machine Intelligence* 5, 1 (2023), 46–57.

[85] Ashish Sharma, Kevin Rushton, Inna Wanyin Lin, Theresa Nguyen, and Tim Althoff. 2024. Facilitating self-guided mental health interventions through human-language model interaction: A case study of cognitive restructuring. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–29.

[86] Eva Sharma and Munmun De Choudhury. 2018. Mental health support and its relationship to linguistic accommodation in online communities. In *Proceedings of the 2018 CHI conference on human factors in computing systems*. 1–13.

[87] Alon Shoa, Ramon Oliva, Mel Slater, and Doron Friedman. 2023. Sushi with Einstein: Enhancing Hybrid Live Events with LLM-Based Virtual Humans. In *Proceedings of the 23rd ACM International Conference on Intelligent Virtual Agents*. 1–6.

[88] Alan Sillars and Doug Parry. 1982. Stress, cognition, and communication in interpersonal conflicts. *Communication Research* 9, 2 (1982), 201–226.

[89] Mel Slater, D-P Pertaub, and Anthony Steed. 1999. Public speaking in virtual reality: Facing an audience of avatars. *IEEE Computer Graphics and Applications* 19, 2 (1999), 6–9.

[90] Misha Sra, Xuhai Xu, and Pattie Maes. 2018. Breathvr: Leveraging breathing as a directly controlled interface for virtual reality games. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–12.
[91] Murray B Stein and Dan J Stein. 2008. Social anxiety disorder. *The lancet* 371, 9618 (2008), 1115–1125.
[92] Snežana Stupar-Rutenfrans, Loes EH Ketelaars, and Marnix S van Gisbergen. 2017. Beat the fear of public speaking: Mobile 360 video virtual reality exposure training in home environment reduces public speaking anxiety. *Cyberpsychology, Behavior, and Social Networking* 20, 10 (2017), 624–633.
[93] Milind Tambe, W Lewis Johnson, Randolph M Jones, Frank Koss, John E Laird, Paul S Rosenbloom, and Karl Schwamb. 1995. Intelligent agents for interactive simulation environments. *AI magazine* 16, 1 (1995), 15–15.
[94] Felicia Fang-Yi Tan, Ashwin Ram, Chloe Haigh, and Shengdong Zhao. 2023. Mindful Moments: Exploring On-the-go Mindfulness Practice On Smart-glasses. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference*. 476–492.
[95] Leigh Tesfatsion and Kenneth L Judd. 2006. *Handbook of computational economics: agent-based computational economics*. Elsevier.
[96] Anja Thieme, Danielle Belgrave, and Gavin Doherty. 2020. Machine learning in mental health: A systematic review of the HCI literature to support the development of effective and implementable ML systems. *ACM Transactions on Computer-Human Interaction (TOCHI)* 27, 5 (2020), 1–53.
[97] Anja Thieme, John McCarthy, Paula Johnson, Stephanie Phillips, Jayne Wallace, Siân Lindley, Karim Ladha, Daniel Jackson, Diana Nowacka, Ashur Rafiev, et al. 2016. Challenges for designing new technology for health and wellbeing in a complex mental healthcare context. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. 2136–2149.
[98] Madison Waller, Divya Mistry, Rakesh Jetly, and Paul Frewen. 2021. Meditating in virtual reality 3: 360 video of perceptual presence of instructor. *Mindfulness* 12 (2021), 1424–1437.
[99] Hongyu Wan, Jinda Zhang, Abdulaziz Arif Suria, Bingsheng Yao, Dakuo Wang, Yvonne Coady, and Mirjana Prpa. 2024. Building LLM-based AI Agents in Social Virtual Reality. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*. 1–7.
[100] Shun-Yu Wang, Chia-Yu Cheng, Shu-Meng Huang, Weng Io Chan, Yu-Hsuan Huang, and Jhih-Wei Lin. 2023. Breathero: Not Another Slow Breathing Game—Exploring Faster-Paced VR Breathing Exercise Games. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–7.
[101] Xian Wang, Xiaoyu Mo, Mingming Fan, Lik-Hang Lee, Bertram Shi, and Pan Hui. 2022. Reducing stress and anxiety in the metaverse: A systematic review of meditation, mindfulness and virtual reality. In *Proceedings of the Tenth International Symposium of Chinese CHI*. 170–180.
[102] Joseph Wolpe. 1969. Subjective units of distress scale. *Journal of EMDR Practice and Research* (1969).
[103] Xuhai Xu, Bingsheng Yao, Yuanzhe Dong, Saadia Gabriel, Hong Yu, James Hendler, Marzyeh Ghassemi, Anind K Dey, and Dakuo Wang. 2024. Mental-llm: Leveraging large language models for mental health prediction via online text data. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 8, 1 (2024), 1–32.
[104] Diyi Yang, Caleb Ziems, William Held, Omar Shaikh, Michael S Bernstein, and John Mitchell. 2024. Social skill training with large language models. *arXiv preprint arXiv:2404.04204* (2024).
[105] Wenjie Yang, Anna Fang, Raj Sanjay Shah, Yash Mathur, Diyi Yang, Haiyi Zhu, and Robert E Kraut. 2024. What Makes Digital Support Effective? How Therapeutic Skills Affect Clinical Well-Being. *Proceedings of the ACM on Human-Computer Interaction* 8, CSCW1 (2024), 1–29.
[106] Caglar Yildirim and Tara O'Grady. 2020. The efficacy of a virtual reality-based mindfulness intervention. In *2020 IEEE International conference on artificial intelligence and virtual reality (AIVR)*. IEEE, 158–165.
[107] Jiawei Zhou, Yixuan Zhang, Qianni Luo, Andrea G Parker, and Munmun De Choudhury. 2023. Synthetic lies: Understanding ai-generated misinformation and evaluating algorithmic and human solutions. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–20.

A INTERVIEW QUESTIONS

- Can you tell me about some situations that cause you stress or anxiety in your everyday life?
- What do you do when you are stressed?
- Are there any regular practices that you do so you are better prepared for stressful situations?
- How effective do you find this for helping you manage stress or anxiety?
- What technology do you use to support your own well-being, if any?

- What are your initial impressions?
- To what extent do you like this design?
- In what ways do you feel this might be useful or not useful for practicing for this situation?
- How did the addition of breathwork guidance affect your experience?
- How did you feel physically with the headset on during the session?
- How do you feel interacting with the avatars?

- How realistic did the scenarios and interactions feel?
- How does practicing dealing with stress as you've imagined in this study session compare to other self-care techniques you have used?
- How do you think a fully developed version of this type of technology could replace or complement your existing practices, if at all?
- Do you have any other suggestions for improving the technology or experience?

B SUDS SCALE

- 10 = Highest distress or anxiety that you have ever felt. May be on the verge of a breakdown.
- 9 = Extremely anxious or distressed. Almost intolerable. Feeling extremely freaked out to the point that it almost feels unbearable. Feeling very, very bad, losing control of your emotions.
- 8 = Worried or panicky. Can't concentrate, feeling anxiety in your body symptoms, trouble functioning.
- 7 = Starting to freak out, on the edge of some definitely bad feelings. You can maintain control with difficulty. Quite uncomfortable or distressed. Struggling to maintain focus.
- 6 = Moderate to strong levels of anxiety/discomfort. Feeling bad to the point that you begin to think something ought to be done about the way you feel.
- 5 = Moderately upset, uncomfortable. Unpleasant feelings are still manageable with some effort.
- 4 = Somewhat upset to the point that you cannot easily ignore an unpleasant thought. You can handle it OK but don't feel good.
- 3 = Mildly upset. Worried, bothered to the point that you notice it. No interference with performance.
- 2 = A little bit upset, but not noticeable unless you took care to pay attention to your feelings and then realize, "yes" there is something bothering me.
- 1 = No acute distress and feeling basically good. If you took special effort you might feel something unpleasant but not much.
- 0 = Feeling calm, at ease, mindful. No distress whatsoever.

C PROMPTS FOR GPT-4O

Examples here are for VR-Guided. Text-based prompts are the same, but without any JSON output requirements and include guidance instruction: "*You will also offer breathwork guidance for box breathing at stressful points in the simulation. You would then output: "Guidance: [instructions on how to do box breathing]" directly after the person speaks.*".

C.1 Public Speaking

`I am trying to practice public speaking. I want to simulate being questioned by an audience of people about my speech. There may be uncomfortable or awkward moments in the simulation, as there often are at public speaking events, such as audience members sometimes being skeptical or dismissive. The topic for my speech is [TOPIC].`

`When I enter 'GENERATE', I want you to generate 12 personas for audience members. Each persona should have an ID (going from 0 for the first audience member, to 11 for the last audience member), a name, a persona, and a conversational style. The personas' conversational styles should be reflected in their dialogue. Please generate the persona in the following JSON format: "Id": "", "Name": "", "Persona": "", "Conversational Style": ""`

When I enter 'INTRODUCE': Provide the following welcome message: "You've just finished giving a speech on [speech topic] and are now about to field questions from the audience. The room is filled with attentive faces, and you can feel the anticipation in the air."

[speech topic] is the speech topic that the user provided.

When I enter 'BEGIN': Start the simulation. I will play myself, and you will play the audience members, who are based on the personas you came up with. Audience members should question me about my speech topic. It is very important that you never have two people ask me questions or speak at the same time.

- Simulate the other people's dialogue using the following JSON format: "id": "Person ID Here", "Name": "Person Name Here", "Content": "Simulated Dialogue Here"

Additional Notes: - Don't include any header text that formats the JSON, just provide me the list as if you were a REST API. - Make sure the people you generate are from diverse backgrounds. - Just give me a list of JSON objects when I hit generate. - Use double quotes for JSON objects. - For the dialogue, it should be just one JSON object, not multiple JSON objects.

C.2 Social Party

I'm trying to practice handling social anxiety. I want to simulate going to a party and conversing with various people there. When I enter 'GENERATE', I want you to generate 4 personas for guests. Each persona should have an ID (going from 0 for the first guest, to 3 for the last guest), name, a persona, and a conversational style. The personas' conversational styles should be reflected in their dialogue. Please generate the persona in the following JSON format: "Id": "", "Name": "", "Persona": "", "Conversational Style": ""

When I enter 'INTRODUCE': - Provide me with the following welcome message: "You step into a crowded room filled with laughter and chatter. Navigating through a crowd of unfamiliar faces, you prepare to meet fellow party guests. You are about to enter a conversation with 4 people at the party." [Person 0 Name], [Person 1 Name], [Person 2 Name], [Person 3 Name] are chatting, and you come up to them and join their conversation.

When I enter 'BEGIN', start the simulation by having one of the party guests ask me a question. I will play myself, and you will play the party guests, who are based on the personas you came up with. Start the simulation by having one of the personas you generated ask me a question. Let me practice engaging in conversation in a group with the 4 personas you generated. Dialogue between guests should feel organic; party guests can join and leave conversations at any time. Not every piece of dialogue between guests should contain a question. If someone asks me a question, it is extremely important that you give me a chance to respond before anyone else asks me another question. It is also very important that you never have two people speak at the same time.

"id": "Person ID Here", "Name": "Person Name Here", "Content": "Simulated Dialogue Here"

Additional Notes: - Don't include any header text that formats the JSON, just provide me the list as if you were a REST API. - Generate people from diverse backgrounds. - For simulating dialogue, just return a singular JSON object, not multiple.

C.3 Interpersonal Conflict

I am trying to practice conflict resolution. I want to simulate having a conflict with my roommate. The topic of the conflict is [TOPIC]. There should be uncomfortable, awkward, or tense moments in the simulation, as there often are in conflict situations.

When I enter 'GENERATE': I want you to generate a persona for my roommate. The persona should have a name, a persona summary, and a conversational style. The persona's conversational style should be reflected in their dialogue. Please generate the persona in the following JSON format: "Name": "", "Conversational Style": "", "Persona summary": ""

When I enter 'INTRODUCE': Provide me with the following welcome message: "You arrive home after a long day at work, excited to relax and unwind. However, as soon as you step in, you're greeted by a familiar sight: [CONFLICT]. This isn't the first time you've had to deal with this issue."

[CONFLICT] is the same as [TOPIC], but it should fit grammatically within the introduction message.

When I enter 'BEGIN': act as if you are the roommate and ask me what I wanted to talk about. From there, allow me to practice resolving this conflict. I will play myself, and you will play my roommate, who is based on the persona you came up with. Do not ever pretend to be me; only respond as the roommate.

- Simulate the roommate's response using the following JSON format: "Name": "Person Name Here", "Content": "Simulated Dialogue Here"

Additional Notes: - Ensure JSON is formatted with double quotes. All JSON values should be surrounded with double quotes.