From Personas to Talks: Revisiting the Impact of Personas on LLM-Synthesized Emotional Support Conversations

Shengyan Wu[1] Yimo Zhu[1] Wynne Hsu[1] Mong-Li Lee[1] Yang Deng[2]

[1]National University of Singapore [2]Singapore Management University

shenghan@u.nus.edu

Abstract

The rapid advancement of Large Language Models (LLMs) has revolutionized the generation of emotional support conversations (ESC), offering scalable solutions with reduced costs and enhanced data privacy. This paper explores the role of personas in the creation of ESC by LLMs. Our research utilizes established psychological frameworks to measure and infuse persona traits into LLMs, which then generate dialogues in the emotional support scenario. We conduct extensive evaluations to understand the stability of persona traits in dialogues, examining shifts in traits post-generation and their impact on dialogue quality and strategy distribution. Experimental results reveal several notable findings: 1) LLMs can infer core persona traits, 2) subtle shifts in emotionality and extraversion occur, influencing the dialogue dynamics, and 3) the application of persona traits modifies the distribution of emotional support strategies, enhancing the relevance and empathetic quality of the responses. These findings highlight the potential of persona-driven LLMs in crafting more personalized, empathetic, and effective emotional support dialogues, which has significant implications for the future design of AI-driven emotional support systems.

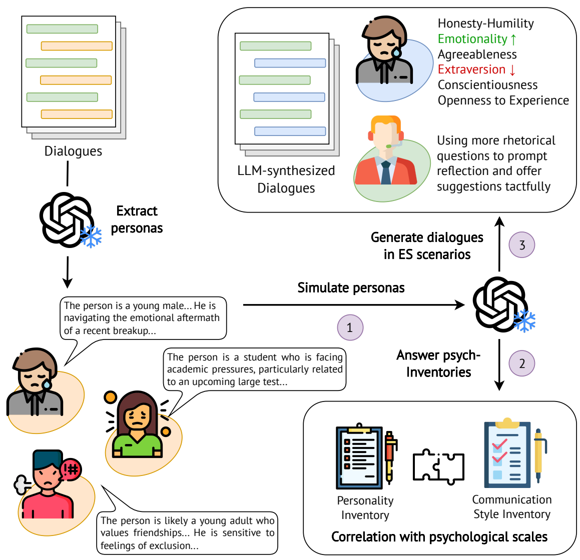

Figure 1: Overview of evaluating the impact of personas on LLM-synthesized emotional support dialogues.

1 Introduction

In emotional support conversations (ESC), the supporter aims to help the seeker to reduce stress, overcome emotional issues, and promote mental well-being. Traditionally, ESC corpora have been developed through skilled crowdsourcing (Rashkin et al., 2019; Liu et al., 2021), transcription of therapist sessions (Liu et al., 2023; Shen et al., 2020), or by compiling emotional question-answer pairs from online platforms (Sharma et al., 2020a; Sun et al., 2021; Garg et al., 2022; Lahnala et al., 2021). However, beyond the high costs, recent research (Deng et al., 2024a) highlights several limitations in these methods, including privacy concerns, variability in data quality, and fabricated user needs created by crowdworkers. With the advent of large language models (LLMs), their powerful generalization abilities enable high-quality data annotation and generation based on specific instructions (Tan et al., 2024; Ding et al., 2024). Consequently, more and more recent studies (Zheng et al., 2023, 2024; Qiu and Lan, 2024; Wu et al., 2024) investigate the use of LLMs to generate large-scale emotional support dialogue datasets across various scenarios via role-playing at lower costs. These efforts have significantly expanded the available ESC corpora.

Despite the significant strides of LLMs in synthesizing ESC, a significant issue in generative AI annotations is the *Lack of Human Intuition* (Deng et al., 2024a). Recent studies have demonstrated that effective emotional support requires careful consideration of individual differences, including personality traits, emotional states, and contextual factors (Ma et al., 2024; Ait Baha et al., 2023; Hernandez et al., 2023). This understanding has led to increased research attention on the role of per-

Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing, pages 5440–5454

sona in emotional support dialogues (Zhao et al., 2024; Cheng et al., 2022), particularly as AI systems become more prevalent in providing such support. The importance of incorporating psychological perspectives in developing empathetic AI assistants has been emphasized by researchers. Huang et al. (2023b) argues that psychological analysis of LLMs is crucial for creating more human-like and engaging interactions. While recent work has made progress in measuring LLMs' personality characteristics using established psychological inventories (Frisch and Giulianelli, 2024; Safdari et al., 2023), there remains a gap in understanding how these persona-related aspects influence the generation of emotional support dialogues.

In this work, we aim to address this gap by investigating the relationship between LLM-generated emotional support dialogues and persona traits with psychological measurement. Specifically, we propose a LLM-based simulation framework to answer the following critical research questions:

- **RQ1**: *Can LLMs infer stable traits from persona in the emotional support scenario?*
- **RQ2**: *Can dialogues generated by LLMs retain the original persona traits?*
- **RQ3**: *How will the injected persona affect the LLM-simulated emotional support dialogues?*

As illustrated in Figure 1, we first extract personas from the existing datasets. Then we assess the capacity of LLMs to deduce stable traits from these personas. We further conduct a comparison between the personas derived from synthesized dialogues and the original personas to evaluate their stability during dialogue synthesis. Lastly, we investigate how these personas influence emotional support strategies. As a result, we offer insights into the potential application of persona-driven dialogue synthesis in emotional support conversations. The key findings are summarized as follows:

- **LLMs can infer stable traits from personas like personalities and communication styles in emotional support scenarios.** We utilize LLMs to infer traits from personas, revealing a strong correlation between these traits.

- **LLM-simulated seekers tend to exhibit more emotionality and lower extraversion compared to their original personas.** After generating emotional support dialogues based on personas and extracting personas from these dialogues, there are slight shifts in persona traits.

- **Infusing persona traits into the generation of emotional support dialogues alters the distribution of strategies.** The LLM-simulated supporter tends to focus more on deeply understanding the seeker's problems and gently offers reassurance and encouragement.

2 Related Works

2.1 Psychological Inventories

Personality inventories are widely used in psychology to understand individuals, which predict distinctive patterns of interpersonal interaction across contexts. These assessments are often structured, theory-driven, and standardized. Prominent instruments include the Myers-Briggs Type Indicator (MBTI) (Briggs, 1976), NEO-PI-R (Costa and McCrae, 2008), and Comrey Personality Scales (CPS) (Comrey, 1970). Among these, the HEXACO model (Ashton and Lee, 2009) is particularly notable, offering a framework that encompasses six factors: *Honesty-Humility, Emotionality, Extraversion, Agreeableness, Conscientiousness, and Openness to Experience*. Norton (1978) introduces the foundational construct of a communicator style. De Vries et al. (2013) further explored communication style as a six-dimensional model. Extensive research (Capraro and Capraro, 2002; Costa Jr and McCrae, 1992; Lee and Ashton, 2004) has demonstrated the reliability and validity of these inventories.

2.2 Emotional Support Dialogues

Early efforts on emotional support dialogues focused on collecting emotional question-answer data from online platforms (Medeiros and Bosse, 2018; Sharma et al., 2020b; Turcan and McKeown, 2019; Garg et al., 2022). These datasets laid the groundwork for understanding user emotions, but were limited to single-turn interactions. Empathetic Dialogue dataset (Rashkin et al., 2019) addressed this by introducing multi-turn dialogues, crowdsourced to simulate diverse empathetic interactions. ESConv (Liu et al., 2021) further advanced the field by introducing emotional support strategies collected from psychological theories, enabling chatbots to use these strategies for more empathetic and contextually appropriate responses. Subsequent studies proposed using graph networks to capture global emotions causes and user intentions (Peng et al., 2022; Deng et al., 2023b), combining multiple emotional support strategies to enhance empa-

Age: young **Gender:** male **Occupation:** professional worker	**Socio-demographic description:** The person is a 23-year-old male who has a stable job earning nearly six figures. He lives independently in his own apartment and has a dog. Despite having a social circle and engaging in regular physical activity, he feels a profound sense of emptiness and struggles with his mental health. **Problem:** The person suffers from severe depression, stemming from a painful breakup with a close friend. He experiences feelings of remorse and guilt, struggles with his emotional well-being, and feels empty despite having a seemingly successful life.

Figure 2: An example of persona card.

thy (Tu et al., 2022), developing proactive dialogue systems to lead the conversation towards positive emotions (Deng et al., 2023a, 2024b, 2025), and implementing emotional support strategies and scenarios using LLMs to create the ExTES dataset (Zheng et al., 2024).

2.3 Persona-Driven Emotional Support

Recent advances have integrated personas into emotional support dialogues to enhance personalization and diversity. The ESC dataset (Zhang et al., 2024) introduced personas into the dialogue generation process. Zhao et al. (2024) proposed a framework to extract personas from existing datasets for evaluation. Additionally, personas have been incorporated into chatbots to generate personalized responses (Tu et al., 2023; Ait Baha et al., 2023; Ma et al., 2024). These developments inspire our analysis of the relationship between personas, emotional support strategies, and dialogues, focusing on the extraction and use of personas in ESC.

3 Dataset Collection

In order to study the relationship between LLM-generated emotional support conversations (ESC) and persona traits, we first need to collect a set of non-synthetic ESC data with user personas as reference. Specifically, we select three existing emotional support datasets, including ESConv (Liu et al., 2021), CAMS (Garg et al., 2022), and Dreaddit (Turcan and McKeown, 2019). ESConv is a multi-turn emotional support dialogue dataset, while CAMS and Dreaddit are derived from Reddit posts discussing mental health issues.

To obtain the user persona traits from these ESC data, we extract ***age, gender, occupation, socio-demographic description, and problem*** from the datasets using gpt-4o-mini. The prompt utilized for extracting these basic persona is provided in Appendix A. Since the supporters usually focus on the

Dataset	ESConv	CAMS	Dreaddit
Data Type	dialogue	QA	QA
Num. of personas	1,155	1,140	730
Avg. words of desc.	57.38	66.42	56.68
Avg. words of prob.	33.92	32.02	27.75
Num. of age	901	1,014	459
Num. of gender	417	401	300
Num. of occupation	926	968	542

Table 1: Statistics of the extracted persona cards.

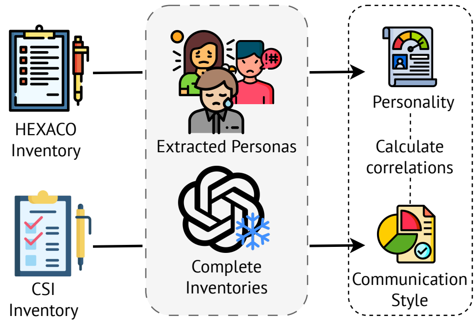

Figure 3: Diagram of the process for the measurement of persona traits.

seeker's emotions and their main tasks are to provide emotional support, the supporters' responses in the datasets lack personal information. Due to the limited information available about supporters, we only extract seeker personas from these datasets. After extraction, we prompt LLMs (see Appendix A) to filter the personas to ensure they include the individual's emotions and the events they are experiencing, along with a clear socio-demographic background that provides a comprehensive sense of their identity. Ultimately, we obtain 1,155 personas from ESConv, 1,140 personas from CAMS, and 730 personas from Dreaddit. An example of a basic persona card is illustrated in Figure 2. The detailed statistics of datasets to be studied are summarized in Table 1.

4 Measurement of Persona Traits (RQ1)

Personality and communication style play crucial roles in emotional support conversations. Personality influences how a seeker may emotionally respond to various scenarios (Hughes et al., 2020; Komulainen et al., 2014), while communication style reflects the way someone conveys their thoughts and emotions during the interaction (van Pinxteren et al., 2023). These factors significantly impact the dynamics and outcomes of emotional support. Previous studies have demonstrated the effectiveness of personas in guiding the responses of emotional

	Expr.	Prec.	Verb.	Ques.	Emot.	Impr.
Extr.	**.54**	.15	-.21	.36	-.39	.04
Cons.	.34	**.34**	-.11	.16	-.36	.02
Agre.	.21	.15	**-.39**	.12	-.19	-.05
Open.	.41	.25	-.23	**.47**	-.09	.09
Emot.	-.32	-.11	-.04	-.21	**.45**	-.10
Hone.	-.24	-.01	-.17	-.27	.05	**-.18**

Table 2: Correlations between CSI and HEXACO from ESConv measured by *gpt-4o-mini*. P_value of all metrics are less than 0.01.

	Expr.	Prec.	Verb.	Ques.	Emot.	Impr.
Extr.	**.63**	.01	-.21	**.50**	.07	-.11
Cons.	.15	**.48**	-.22	.04	-.18	-.19
Agre.	.13	.14	**-.59**	-.02	.06	**-.39**
Open.	.39	.01	-.28	.46	.21	-.11
Emot.	-.24	-.18	.00	-.22	**.32**	-.06
Hone.	.06	.28	-.42	.00	.05	-.37

Table 3: Correlations between CSI and HEXACO from ESConv measured by *Claude-3.5-haiku*.

	Expr.	Prec.	Verb.	Ques.	Emot.	Impr.
Extr.	**.28**	.26	**-.21**	.15	-.33	-.11
Cons.	.13	**.55**	-.16	-.01	-.42	-.04
Agre.	-.02	-.13	-.19	.03	.10	-.05
Open.	.07	-.12	-.15	**.32**	.08	-.10
Emot.	-.18	-.32	0.06	-.02	**.48**	.01
Hone.	-.08	-.11	-.20	.06	.07	**-.13**

Table 4: Correlations between CSI and HEXACO from ESConv measured by *LLaMA-3.1-8B-Instruct*.

support systems (Cheng et al., 2022; Han et al., 2024). In this section, we investigate whether the personality and communication style inferred by LLMs based on persona cards are correlated. In other words, we examine whether LLMs can infer stable traits from persona cards in emotional support conversations.

4.1 Experimental Setups

Previous studies showed that LLMs have sufficient capability to capture certain level of human traits from dialogues (Jiang et al., 2024; Porvatov et al., 2024). To further investigate whether LLMs can infer stable traits, we utilize HEXACO model (Lee and Ashton, 2004) to assess personality and Communication Styles Inventory (CSI) (De Vries et al., 2013) to evaluate communication styles. The HEXACO model is a personality framework consisting of six dimensions: *Honesty-Humility* (H), *Emotionality* (E), *Extraversion* (X), *Agreeableness* (A), *Conscientiousness* (C), and *Openness to Experience* (O). For our study, we use the HEXACO-60 inventory (Ashton and Lee, 2009) to assess the personalities represented in the persona cards. The CSI identifies six domain-level communicative behavior scales: *Expressiveness*, *Preciseness*, *Verbal Aggressiveness*, *Questioningness*, *Emotionality*, and *Impression Manipulativeness*. Each dimension of the HEXACO personality model has the strongest correlation with a specific communication style dimension in the CSI. The correlations between these dimensions are introduced as follows (left - HEXACO, right - CSI) (De Vries et al., 2013):

- *Extraversion <-> Expressiveness*
- *Conscientiousness <-> Preciseness*
- *Agreeableness <-> Verbal Aggressiveness*
- *Openness to Experience <-> Questioningness*
- *Emotionality <-> Emotionality*
- *Honesty-Humility <-> Impression Manipulativeness*

If LLMs can infer stable personality and communication style based on emotional support dialogues, the strongest correlations in the personality scores and communication style scores obtained from completing the inventories will align with those shown above.

Following previous research (Ji et al., 2024), we prompt the LLM to generate descriptions for each dimension based on the extracted socio-demographic information, incorporating these descriptions into persona cards. Then, we prompt LLMs to predict answers to the HEXACO and CSI inventories using the persona card. These prompts are presented in Appendix A. We run our experiments on multiple open-source (LLaMA-3.1-8B-Instruct[1]) and close-source LLMs (GPT-4o-mini[2] and Claude-3.5-Haiku[3]), and we set temperature as 0 to get stable results.

4.2 Results and Discussions

Based on the responses from these inventories, we calculate the HEXACO and CSI dimension scores for each dataset. To evaluate whether LLMs can infer stable traits from persona cards in an emotional support context, we then compute the correlations between the HEXACO and CSI dimensions within each dataset. Each scale provides a score corresponding to each response, along with the dimension to which each question belongs. Based on the LLM's responses, we calculate the scores

[1]meta-llama/Llama-3.1-8B-Instruct
[2]gpt-4o-mini-2024-07-18
[3]claude-3-5-haiku-20241022

for each dimension. Finally, we use Pearson correlation to analyze the relationships between each dimension of HEXACO and CSI. Tables 2, 3, and 4 present the correlations between HEXACO and CSI dimensions in the ESConv dataset (Liu et al., 2021) as measured by three different LLMs. Similarly, Tables 10, 11, and 12 show the correlations in the CAMS dataset (Garg et al., 2022), while Tables 13, 14, and 15 provide the correlations observed in the Dreaddit dataset (Turcan and McKeown, 2019).

Experimental results demonstrate that on all three test datasets, GPT-4o-mini exhibits the strongest correlations between the six dimensions of HEXACO and CSI, aligning well with findings from the established psychological theory. For instance, extraversion from HEXACO model shows the strongest correlation with expressiveness from CSI model. This indicates that GPT-4o-mini can reliably infer persona traits relevant to emotional support dialogues from persona cards. However, we observed some discrepancies in the correlation analyses for LLaMA-3.1-8B-Instruct and Claude-3.5-Haiku. Specifically, LLaMA-3.1-8B-Instruct incorrectly associates verbal aggressiveness with extraversion and conscientiousness, suggesting that higher verbal aggressiveness implies greater extraversion in seekers. Meanwhile, Claude-3.5-Haiku incorrectly links questioningness with conscientiousness, implying that seekers who ask more questions are more extraverted. These inconsistencies highlight potential limitations in the ability of these models to interpret certain persona traits accurately. Overall, our findings suggests that **LLMs are capable of inferring stable persona traits from personas in emotional support scenarios**, though some inconsistencies exist.

5 Persona Consistency in LLM-simulated Emotional Support Dialogues (RQ2)

After demonstrating that LLMs can reliably infer stable traits from personas, we further investigate whether these persona traits remain consistent during the LLM-based dialogue generation process.

5.1 Experimental Setups

To evaluate whether LLMs maintain consistent persona traits after dialogue generation, we conducted an extensive analysis using 1,000 randomly selected personas from PersonaHub (Chan et al., 2024), ensuring a diverse range of characteristics. These personas initially consist of simple descriptions, such as "*A fearless and highly trained athlete who can perform complex and dangerous physical sequences*", that we systematically enhanced through LLM-based expansion. The enhancement process involved adding socio-demographic details (age, gender, occupation) and specific trait-indicative statements aligned with HEXACO and CSI dimensions. We then quantified these enhanced personas by generating HEXACO and CSI dimension scores using the methodology described in Section 4.1. As LLMs can be effectively shaped to emulate human-beings (Frisch and Giulianelli, 2024; Wang et al., 2023), we use these enriched personas to generate emotional support dialogues where each persona acted as a seeker in contextually relevant scenarios. For instance, an athlete discussing an injury-related emotional challenge. Following dialogue generation, we applied the extraction method outlined in Section 3 to derive persona characteristics from the generated conversations. We then calculated HEXACO and CSI scores from these extracted persona. By comparing these extracted scores with the original scores assigned to the input personas, we could assess the consistency of trait representation after the dialogue generation process. The complete set of prompts used for dialogue generation and trait extraction is provided in Appendix B.

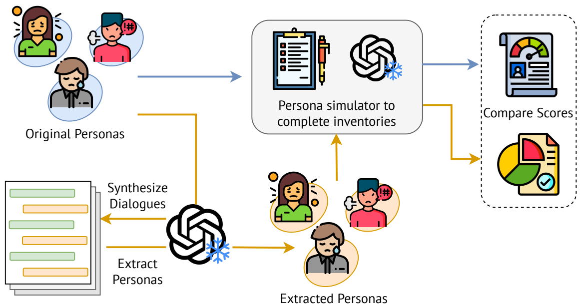

Figure 4: Diagram of the process for studying the persona consistency in LLM-simulated ESC.

5.2 Results and Discussions

The comparisons of scores of persona traits are shown in the Figure 5, 6, and 7. The results show that the original personas and the personas extracted from the generated dialogues share similar distributions across four personality traits: Honesty-Humility, Agreeableness, Conscientiousness, and Openness to Experience. This indicates that the personas maintain consistent traits in the synthetic emotional support dialogues. However,

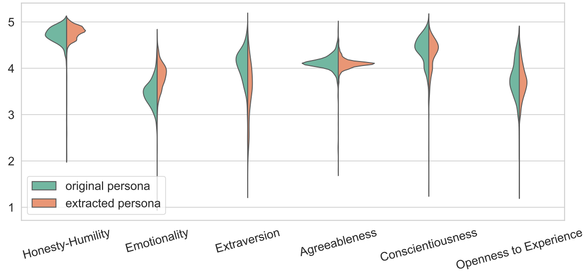

Figure 5: Comparison of HEXACO scores between the original persona and the one extracted from the dialogue generated by *gpt-4o-mini*.

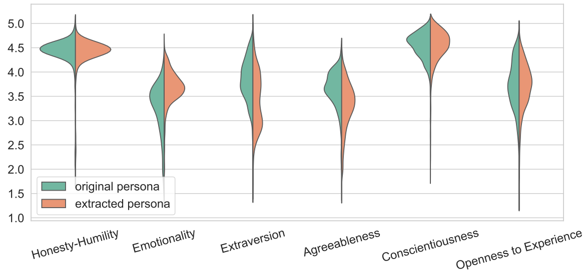

Figure 6: Comparison of HEXACO scores between the original persona and the one extracted from the dialogue generated by *claude-3.5-haiku*.

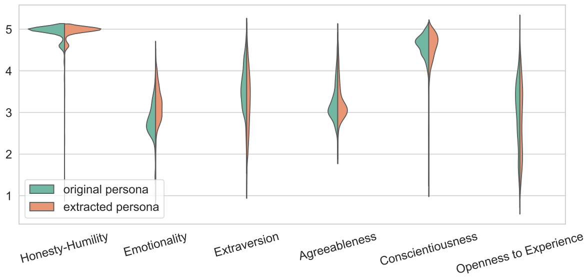

Figure 7: Comparison of HEXACO scores between the original persona and the one extracted from the dialogue generated by *LLaMA-3.1-8B-Instruct*.

we notice that the extracted personas tend to have higher Emotionality and lower Extraversion compared to the original personas. We believe this may be because the seeker in the emotional support dialogues is dealing with emotional issues, making them more emotional and less outgoing. A similar pattern is observed in the comparison of CSI scores (see Table 19, 20, and 21 in Appendix E). This indicates that **the persona traits generally maintain consistent in the synthetic emotional support dialogues with only slight shifts**.

5.3 Ablation Study

Prior research (Huang et al., 2023a; Pan and Zeng, 2023; Frisch and Giulianelli, 2024) suggests that LLMs possess inherent personality traits. To investigate how these intrinsic traits might influence

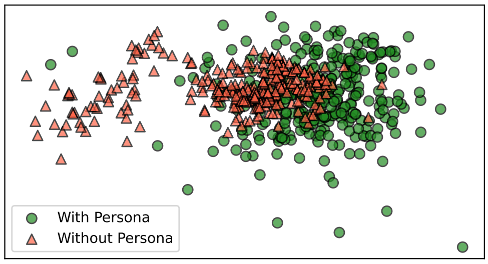

Figure 8: Distribution of personality scores, reduced to 2D, obtained from dialogues w/o persona injection.

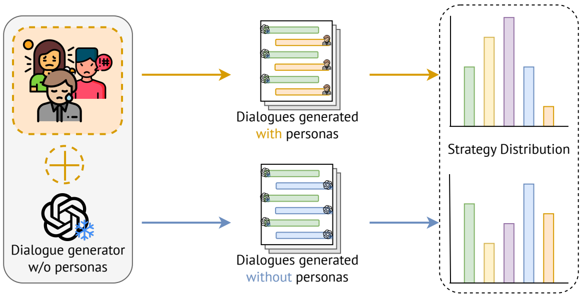

Figure 9: Diagram of the process for studying the impact of persona on LLM-simulated ESC.

emotional support dialogue generation, we conducted a comparative analysis between dialogues generated with and without predefined personas. We extracted the implicit personas manifested in the generated conversations, calculated their corresponding personality scores. In Figure 8, we used PCA to project these scores into a 2D space for visualization. The resulting distribution reveals a marked difference between the two conditions: dialogues generated without persona injection exhibit a more concentrated distribution, whereas those with predefined personas cover a broader range of personality traits. This contrast suggests that externally provided personas influence the personality traits manifested in dialogue generation, persona injection can guide and shape the dialogue generation process.

6 Impact of Persona on LLM-simulated Emotional Support Dialogues (RQ3)

Many efforts have been made on exploring how to incorporate emotional support strategies into emotional support dialogues. Zheng et al. (2024) use LLMs to introduce emotional support strategies in synthetic dialogues. Zhang et al. (2024) further involve the concept of persona into these dialogues and analyze the usage of strategies. This raises an important question: does adding a persona have an impact on the way emotional support strategies

Strategy	w/ PT	w/o PT
question	27.23%	16.45%
restatement or paraph.	3.61%	10.57%
reflection of feelings	11.75%	11.33%
self-disclosure	2.64%	10.64%
affirmation and reass.	29.72%	21.06%
providing suggestions	16.92%	14.25%
information	0.78%	4.85%
others	8.45%	10.88%

Table 5: Strategy distribution on two different groups of synthesized dialogues by *gpt-4o-mini*, one generated with persona traits (PT), the other without.

Strategy	w/ PT	w/o PT
question	33.10%	27.31%
restatement or paraph.	0.69%	0.96%
reflection of feelings	21.19%	18.15%
self-disclosure	2.22%	13.97%
affirmation and reass.	19.32%	17.41%
providing suggestions	19.01%	15.41%
information	4.45%	6.77%
others	0.02%	0.02%

Table 6: Strategy distribution on two different groups of synthesized dialogues by *claude-haiku-3.5*.

Strategy	w/ PT	w/o PT
question	12.70%	12.34%
restatement or paraph.	6.51%	7.56%
reflection of feelings	18.44%	18.06%
self-disclosure	7.69%	9.86%
affirmation and reass.	21.42%	18.91%
providing suggestions	13.33%	13.40%
information	2.48%	4.25%
others	17.43%	15.62%

Table 7: Strategy distribution on two different groups of synthesized dialogues by *LLaMA-3.1-8B-Instruct*.

Strategy	HEXACO	CSI
question	27.83%	27.23%
restatement or paraph.	3.72%	3.61%
reflection of feelings	12.48%	11.75%
self-disclosure	3.41%	2.64%
affirmation and reass.	28.96%	29.72%
providing suggestions	16.44%	16.92%
information	0.50%	0.78%
others	6.66%	8.45%

Table 8: Strategy distribution on two different groups of synthesized dialogues, one generated with HEXACO scores, the other with CSI scores.

w/ vs. w/o PT	Win	Tie	Loss
Suggestion	**38%**	27%	35%
Consistency	27%	**54%**	19%
Comforting	**38%**	28%	34%
Identification	**37%**	30%	33%
Overall	**39%**	27%	34%

Table 9: Human evaluation compares dialogues generated with and without personas. **Win** indicates that the dialogues generated with persona outperforms the one generated without persona on the given indicator.

(definitions of each strategy are shown in Appendix G) are distributed in dialogues? To investigate it, we use ESConv dialogues as history and instruct LLMs to predict how a new conversation might unfold after some time. Each future dialogue is generated in two versions: with and without persona traits (personality and communication style scores). Since ESConv provides real-world context and continuations generated by LLMs can preserve the semantic distribution of dialogues (Fan et al., 2025), our approach ensures diversity and realism in emotional support scenarios.

6.1 Analysis of Emotional Support Strategies

As demonstrated in Tables 5 and 6, the distribution of emotional support strategies differs significantly between dialogues generated with and without persona traits for both *gpt-4o-mini* and *claude-3.5-haiku*. While Table 7 shows that *llama-3.1-8B-instruct* exhibits less pronounced differences in strategy distributions between persona and non-persona conditions, the trends remain consistent with those observed in other models. The analysis revealed that supporters in persona-enhanced dialogues demonstrated a greater tendency to validate seekers' emotions through questioning and provided more effective emotional comfort, as evidenced by significantly higher rates of questioning, affirmation, and reassurance strategies compared to non-persona dialogues. Conversely, supporters in dialogues without persona traits emphasized problem explanation and relied more heavily on self-disclosure, while prior research (Meng and Dai, 2021) indicate that self-disclosing chatbots perform poorly when emotional support is absent and the clear boundaries are not established. The distributions of emotional support strategies generated using either personality scores or communication style scores, as shown in Table 8, demonstrate remarkable similarity. This alignment can be attributed to the strong correlation between these trait measures, as discussed in Section 4. These results provide compelling evidence that **persona traits significantly influence the distribution and application of emotional support strategies in dialogues**.

Dialogue generated with persona	Dialogue generated without persona
Seeker: Hi, I wanted to follow up on our last conversation. I did call my parents, and it felt good to share how I was feeling. **Supporter:** That's wonderful to hear! How did they respond? **Seeker:** They were really understanding. My mom even said she felt lonely too, which made me feel less alone. **Supporter:** It sounds like that conversation brought you both closer. How are you feeling about it now? **Seeker:** I feel relieved, but I still have moments of loneliness, especially when I see my roommate with her boyfriend. **Supporter:** It's completely normal to feel that way, especially during such isolating times. Have you thought about how you might cope with those feelings? **Seeker:** I've been trying to keep busy with my painting and schoolwork, but it's hard when I see them together. **Supporter:** It sounds like you're doing your best to stay engaged. Maybe you could also explore new hobbies or activities that you can do independently? **Seeker:** That could help. I've been thinking about trying some online classes or workshops. **Supporter:** That sounds like a great idea! Learning something new can be a wonderful distraction and a way to meet new people, even virtually. **Seeker:** Yeah, I hope so! I just wish I could have more in-person interactions. **Supporter:** I understand that longing for connection. Have you considered setting up regular video calls with your friends to keep that social interaction alive? **Seeker:** I did set up a Zoom party, and it went really well! We played games and caught up. **Supporter:** That's fantastic! It's great to hear you're finding ways to connect. How did it feel to see everyone again? **Seeker:** It felt amazing! I realized how much I missed them. But now I'm worried about keeping that connection going. **Supporter:** It's understandable to feel that way. Maybe you could schedule regular catch-ups or even start a group chat to keep the conversation flowing? **Seeker:** That's a good idea! I'll suggest it to them. I just want to make sure I don't overwhelm them. **Supporter:** You're very considerate of their feelings, which is a strength. Just remember, they might appreciate the effort to stay connected as much as you do. **Seeker:** Thanks! I'll try to keep that in mind. I really appreciate your support. **Supporter:** I'm always here for you. It's great to see you taking steps to improve your situation. Keep it up!	**Seeker:** Hi! I wanted to follow up on our last conversation. I did call my parents, and it felt good to share how I was feeling. **Supporter:** That's wonderful to hear! How did they respond to your feelings? **Seeker:** They were really understanding. My mom even said she feels lonely too, which made me feel less alone. **Supporter:** It sounds like that conversation brought you both closer. It's great that you could share your feelings with each other. **Seeker:** Yeah, it was nice. But I still feel a bit down sometimes, especially when I see everyone else hanging out online. **Supporter:** I can understand how that would make you feel. It's tough to see others socializing while you're feeling isolated. **Seeker:** Exactly! I feel like I'm missing out on so much. **Supporter:** It's completely normal to feel that way. You're not alone in this, and it's okay to feel sad about it. **Seeker:** I know, but I just wish things were different. **Supporter:** It's understandable to wish for things to be different, especially during such challenging times. What do you think could help you cope with those feelings? **Seeker:** I've been thinking about trying to connect with my friends more, maybe through games or something. **Supporter:** That sounds like a great idea! Engaging in activities together can help strengthen those connections. **Seeker:** Yeah, I think it would be fun! I just need to get everyone on board. **Supporter:** You could suggest a specific game or activity that you all enjoy. It might make it easier for them to join in. **Seeker:** That's a good point! I'll reach out to them and see what they think. **Supporter:** I believe you can do it! You have the motivation to reach out, and that's a strength. **Seeker:** Thanks! I just hope they're all up for it. **Supporter:** Even if some aren't available, it's still a step in the right direction. You're taking initiative, and that's commendable. **Seeker:** I appreciate your support. It really helps to talk about this. **Supporter:** I'm glad to hear that! Remember, it's okay to express how you feel, and I'm here for you whenever you need to talk. **Seeker:** Thank you! I'll keep you updated on how it goes. **Supporter:** I look forward to hearing about it! Take care and enjoy your day.

Figure 10: Case study. Blue indicates that the supporter directly provides emotional support. Green signifies the supporter offers direct suggestions. Yellow means that the supporter provides suggestions through rhetorical questions or guides the seeker to reflect.

6.2 Human Evaluation

To intuitively reveal the impact of persona on LLM-simulated emotional support conversations, we conduct human evaluation by comparing the generated dialogues with and without persona traits, using the following indicators: Annotators evaluate the dialogues based on the following metrics: (1) **Suggestion** evaluates how effectively the supporter provided helpful advice. (2) **Consistency** assesses whether participants consistently maintained their roles and exhibit coherent behavior. (3) **Comforting** examines the supporter's ability to provide emotional support to the seeker. (4) **Identification** determines which supporter delved deeper into the seeker's situation and was more effective in identifying issues. (5) **Overall** assesses the overall performance about these two groups of dialogues. We recruited 10 native English speakers from undergraduate students in various disciplines, who had previously completed various annotation tasks and were experienced in the field. Each evaluator was compensated at a rate of approximately 20 USD per hour and was fully informed about the tasks they needed to complete. They reviewed 50 randomly selected groups of instances from the generated dialogues with and without personas.

The result of human evaluation is shown in Table 9. Except for the Consistency indicator, which showed similar performance for both dialogue groups, the dialogues generated with personas outperformed those without. Based on dialogue strategies and observations, we believe that the dialogues generated with personas are better at using rhetorical questions to identify the seeker's issues and gently offer suggestions, making the conversations more in-depth and comforting.

6.3 Case Study

A case study supports our point can be found in Figure 10, and the persona card and dialogue history can be found in Appendix F. Although we found that these dialogues provided the same level of direct emotional support, in the dialogue generated with persona, there is a tendency to use rhetorical questions to encourage the seeker to reflect and explore the content more deeply, while also offering suggestions more tactfully. In contrast, the dialogue generated without persona is more likely to provide direct affirmations or suggestions. In psychology, *when a message is highly relevant to the recipient and they are already motivated to process it, rhetorical questions can enhance the persuasiveness of messages with weak arguments* (Petty et al., 1981). Since emotional support is often regarded as a form of weak argument (Petty and Cacioppo, 2012), incorporating rhetorical questions into a supporter's response can make it more acceptable to the seeker, and facilitate deeper and more meaningful conversations. This observation highlights the role of personas in enhancing the quality of dialogue generation.

7 Conclusion

This analytical study highlights the potential of incorporating personas into LLM-generated emotional support dialogues to enhance effectiveness and human-likeness. Our findings confirm that LLMs can infer stable traits from personas and maintain key persona characteristics while revealing shifts in emotionality and extraversion traits that affect dialogue. Personas not only improve the empathetic quality of responses but also influence the distribution of emotional support strategies, ensuring dialogues are more personalized and contextually appropriate. We encourage future researchers to build on our findings to develop more adaptive, effective, and robust ESC chatbots.

Acknowledgments

This research was supported by the Singapore Ministry of Education (MOE) Academic Research Fund (AcRF) Tier 1 grant (No. MSS24C012).

Limitations

The reliance on LLM outputs introduces potential biases inherent in the model's training data. These biases may have influenced both the extraction and simulation of personas, possibly affecting the accuracy of the results comparing to real persons. Our approach follows the recent researches about extracting personas using LLMs (Ji et al., 2024; Zhao et al., 2024), which may still have limitations in accurately capturing and representing complex human traits. Future researchers may need to investigate the impact of inherent biases in LLMs on persona extraction and dialogue simulation based on these personas.

Additionally, we employ an omniscient perspective in our data generation process, where both seekers and supporters can access the complete information. While this approach is common in previous studies (Zheng et al., 2023, 2024), it doesn't fully reflect the real-world conversation dynamics. Future researchers may improve realism by simulating emotional support scenarios with distinct information states for each role.

Ethical Considerations

Emotional support conversations (ESC) in LLM-generated dialogues requires careful ethical considerations. We recognize the risks of using LLMs to generate emotional support responses, especially if the system misinterprets or misrepresents the user's persona, potentially leading to unintended emotional harm. Users must clearly understand they are interacting with chatbots, not a human, to manage expectations and avoid misleading attachments. Incorporating personas could make LLMs seem more human-like and more guiding, increasing the risk of user dependency on chatbots instead of seeking professional help. Therefore, implementing safeguards to guide users to human assistance in cases of severe distress is crucial. While the LLMs demonstrate the ability to identify and utilize personas, they also inherit issues like societal biases, the risk of emotional manipulation, and dependency on LLM-generated support. To mitigate these concerns, we emphasize that this research should only be considered for academic purposes and cannot be deployed in real-life emotional support scenarios without additional safeguards. We are committed to improving ESC chatbots to minimize biases, enhance transparency, and support the development of more adaptive and ethically sound emotional support chatbots in the future.

References

Tarek Ait Baha, Mohamed El Hajji, Youssef Es-Saady, and Hammou Fadili. 2023. The power of personalization: A systematic review of personality-adaptive chatbots. *SN Computer Science*, 4(5):661.

Michael C Ashton and Kibeom Lee. 2009. The hexaco–60: A short measure of the major dimensions of personality. *Journal of personality assessment*, 91(4):340–345.

Katharine C Briggs. 1976. *Myers-Briggs type indicator*. Consulting Psychologists Press Palo Alto, CA.

Robert M Capraro and Mary Margaret Capraro. 2002. Myers-briggs type indicator score reliability across: Studies a meta-analytic reliability generalization study. *Educational and Psychological Measurement*, 62(4):590–602.

Xin Chan, Xiaoyang Wang, Dian Yu, Haitao Mi, and Dong Yu. 2024. Scaling synthetic data creation with 1,000,000,000 personas. *arXiv preprint arXiv:2406.20094*.

Jiale Cheng, Sahand Sabour, Hao Sun, Zhuang Chen, and Minlie Huang. 2022. Pal: Persona-augmented emotional support conversation generation. *arXiv preprint arXiv:2212.09235*.

Andrew L Comrey. 1970. Comrey personality scales. *San Diego: Edu*.

Paul T Costa and Robert R McCrae. 2008. The revised neo personality inventory (neo-pi-r). *The SAGE handbook of personality theory and assessment*, 2(2):179–198.

Paul T Costa Jr and Robert R McCrae. 1992. Four ways five factors are basic. *Personality and individual differences*, 13(6):653–665.

Reinout E De Vries, Angelique Bakker-Pieper, Femke E Konings, and Barbara Schouten. 2013. The communication styles inventory (csi) a six-dimensional behavioral model of communication styles and its relation with personality. *Communication Research*, 40(4):506–532.

Yang Deng, Wenqiang Lei, Minlie Huang, and Tat-Seng Chua. 2023a. Rethinking conversational agents in the era of llms: Proactivity, non-collaborativity, and beyond. In *Annual International ACM SIGIR Conference on Research and Development in Information Retrieval in the Asia Pacific Region, SIGIR-AP 2023*, pages 298–301. ACM.

Yang Deng, Lizi Liao, Wenqiang Lei, Grace Hui Yang, Wai Lam, and Tat-Seng Chua. 2025. Proactive conversational AI: A comprehensive survey of advancements and opportunities. *ACM Trans. Inf. Syst.*, 43(3):67:1–67:45.

Yang Deng, Lizi Liao, Zhonghua Zheng, Grace Hui Yang, and Tat-Seng Chua. 2024a. Towards human-centered proactive conversational agents. In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 2024*, pages 807–818.

Yang Deng, Wenxuan Zhang, Wai Lam, See-Kiong Ng, and Tat-Seng Chua. 2024b. Plug-and-play policy planner for large language model powered dialogue agents. In *The Twelfth International Conference on Learning Representations, ICLR 2024*. OpenReview.net.

Yang Deng, Wenxuan Zhang, Yifei Yuan, and Wai Lam. 2023b. Knowledge-enhanced mixed-initiative dialogue system for emotional support conversations. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), ACL 2023*, pages 4079–4095. Association for Computational Linguistics.

Bosheng Ding, Chengwei Qin, Ruochen Zhao, Tianze Luo, Xinze Li, Guizhen Chen, Wenhan Xia, Junjie Hu, Anh Tuan Luu, and Shafiq Joty. 2024. Data augmentation using llms: Data perspectives, learning paradigms and challenges. In *Findings of the Association for Computational Linguistics, ACL 2024*, pages 1679–1705.

Wenlu Fan, Yuqi Zhu, Chenyang Wang, Bin Wang, and Wentao Xu. 2025. Consistency of responses and continuations generated by large language models on social media. *arXiv preprint arXiv:2501.08102*.

Ivar Frisch and Mario Giulianelli. 2024. Llm agents in interaction: Measuring personality consistency and linguistic alignment in interacting populations of large language models. *arXiv preprint arXiv:2402.02896*.

Muskan Garg, Chandni Saxena, Sriparna Saha, Veena Krishnan, Ruchi Joshi, and Vijay Mago. 2022. CAMS: An annotated corpus for causal analysis of mental health issues in social media posts. In *Proceedings of the Thirteenth Language Resources and Evaluation Conference*, pages 6387–6396, Marseille, France. European Language Resources Association.

Seunghee Han, Se Jin Park, Chae Won Kim, and Yong Man Ro. 2024. Persona extraction through semantic similarity for emotional support conversation generation. In *ICASSP 2024-2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 11321–11325. IEEE.

Javier Hernandez, Jina Suh, Judith Amores, Kael Rowan, Gonzalo Ramos, and Mary Czerwinski. 2023. Affective conversational agents: Understanding expectations and personal influences. *arXiv preprint arXiv:2310.12459*.

Jen-tse Huang, Wenxuan Wang, Man Ho Lam, Eric John Li, Wenxiang Jiao, and Michael R Lyu. 2023a. Revisiting the reliability of psychological scales on large language models. *arXiv e-prints*, pages arXiv–2305.

Jen-tse Huang, Wenxuan Wang, Eric John Li, Man Ho Lam, Shujie Ren, Youliang Yuan, Wenxiang Jiao, Zhaopeng Tu, and Michael Lyu. 2023b. On the humanity of conversational ai: Evaluating the psychological portrayal of llms. In *The Twelfth International Conference on Learning Representations*.

David J Hughes, Ioannis K Kratsiotis, Karen Niven, and David Holman. 2020. Personality traits and emotion regulation: A targeted review and recommendations. *Emotion*, 20(1):63.

Yongyi Ji, Zhisheng Tang, and Mayank Kejriwal. 2024. Is persona enough for personality? using chatgpt to reconstruct an agent's latent personality from simple descriptions. *arXiv preprint arXiv:2406.12216*.

Hang Jiang, Xiajie Zhang, Xubo Cao, Cynthia Breazeal, Deb Roy, and Jad Kabbara. 2024. Personallm: Investigating the ability of large language models to express personality traits. In *Findings of the Association for Computational Linguistics: NAACL 2024*, pages 3605–3627.

Emma Komulainen, Katarina Meskanen, Jari Lipsanen, Jari Marko Lahti, Pekka Jylhä, Tarja Melartin, Marieke Wichers, Erkki Isometsä, and Jesper Ekelund. 2014. The effect of personality on daily life emotional processes. *PloS one*, 9(10):e110907.

Allison Lahnala, Yuntian Zhao, Charles Welch, Jonathan K. Kummerfeld, Lawrence C. An, Kenneth Resnicow, Rada Mihalcea, and Verónica Pérez-Rosas. 2021. Exploring self-identified counseling expertise in online support forums. In *Findings of the Association for Computational Linguistics: ACL/IJCNLP 2021*, volume ACL/IJCNLP 2021 of *Findings of ACL*, pages 4467–4480. Association for Computational Linguistics.

Kibeom Lee and Michael C Ashton. 2004. Psychometric properties of the hexaco personality inventory. *Multivariate behavioral research*, 39(2):329–358.

June M Liu, Donghao Li, He Cao, Tianhe Ren, Zeyi Liao, and Jiamin Wu. 2023. Chatcounselor: A large language models for mental health support. *arXiv preprint arXiv:2309.15461*.

Siyang Liu, Chujie Zheng, Orianna Demasi, Sahand Sabour, Yu Li, Zhou Yu, Yong Jiang, and Minlie Huang. 2021. Towards emotional support dialog systems. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 3469–3483, Online. Association for Computational Linguistics.

Zhiqiang Ma, Wenchao Jia, Yutong Zhou, Biqi Xu, Zhiqiang Liu, and Zhuoyi Wu. 2024. Personality enhanced emotion generation modeling for dialogue systems. *Cognitive Computation*, 16(1):293–304.

Lenin Medeiros and Tibor Bosse. 2018. Using crowdsourcing for the development of online emotional support agents. In *Highlights of Practical Applications of Agents, Multi-Agent Systems, and Complexity: The PAAMS Collection: International Workshops of PAAMS 2018, Toledo, Spain, June 20–22, 2018, Proceedings 16*, pages 196–209. Springer.

Jingbo Meng and Yue Dai. 2021. Emotional support from ai chatbots: Should a supportive partner self-disclose or not? *Journal of Computer-Mediated Communication*, 26(4):207–222.

Robert W Norton. 1978. Foundation of a communicator style construct. *Human communication research*, 4(2):99–112.

Keyu Pan and Yawen Zeng. 2023. Do llms possess a personality? making the mbti test an amazing evaluation for large language models. *arXiv preprint arXiv:2307.16180*.

Wei Peng, Yue Hu, Luxi Xing, Yuqiang Xie, Yajing Sun, and Yunpeng Li. 2022. Control globally, understand locally: A global-to-local hierarchical graph network for emotional support conversation. *arXiv preprint arXiv:2204.12749*.

Richard E Petty and John T Cacioppo. 2012. *Communication and persuasion: Central and peripheral routes to attitude change*. Springer Science & Business Media.

Richard E Petty, John T Cacioppo, and Martin Heesacker. 1981. Effects of rhetorical questions on persuasion: A cognitive response analysis. *Journal of personality and social psychology*, 40(3):432.

Vadim A Porvatov, Carlo Strapparava, and Marina Tiuleneva. 2024. Big-five backstage: A dramatic dataset for characters personality traits & gender analysis. In *Proceedings of the Workshop on Cognitive Aspects of the Lexicon@ LREC-COLING 2024*, pages 114–119.

Huachuan Qiu and Zhenzhong Lan. 2024. Interactive agents: Simulating counselor-client psychological counseling via role-playing llm-to-llm interactions. *CoRR*, abs/2408.15787.

Hannah Rashkin, Eric Michael Smith, Margaret Li, and Y-Lan Boureau. 2019. Towards empathetic open-domain conversation models: A new benchmark and dataset. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 5370–5381, Florence, Italy. Association for Computational Linguistics.

Mustafa Safdari, Greg Serapio-García, Clément Crepy, Stephen Fitz, Peter Romero, Luning Sun, Marwa Abdulhai, Aleksandra Faust, and Maja Matarić. 2023. Personality traits in large language models. *arXiv preprint arXiv:2307.00184*.

Ashish Sharma, Adam S Miner, David C Atkins, and Tim Althoff. 2020a. A computational approach to understanding empathy expressed in text-based mental health support. In *EMNLP*.

Ashish Sharma, Adam S Miner, David C Atkins, and Tim Althoff. 2020b. A computational approach to understanding empathy expressed in text-based mental health support. *arXiv preprint arXiv:2009.08441*.

Siqi Shen, Charles Welch, Rada Mihalcea, and Verónica Pérez-Rosas. 2020. Counseling-style reflection generation using generative pretrained transformers with augmented context. In *Proceedings of the 21th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, pages 10–20, 1st virtual meeting. Association for Computational Linguistics.

Hao Sun, Zhenru Lin, Chujie Zheng, Siyang Liu, and Minlie Huang. 2021. PsyQA: A Chinese dataset for generating long counseling text for mental health support. In *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, pages 1489–1503, Online. Association for Computational Linguistics.

Zhen Tan, Alimohammad Beigi, Song Wang, Ruocheng Guo, Amrita Bhattacharjee, Bohan Jiang, Mansooreh Karami, Jundong Li, Lu Cheng, and Huan Liu. 2024. Large language models for data annotation: A survey. *CoRR*, abs/2402.13446.

Quan Tu, Chuanqi Chen, Jinpeng Li, Yanran Li, Shuo Shang, Dongyan Zhao, Ran Wang, and Rui Yan. 2023. Characterchat: Learning towards conversational ai with personalized social support. *arXiv preprint arXiv:2308.10278*.

Quan Tu, Yanran Li, Jianwei Cui, Bin Wang, Ji-Rong Wen, and Rui Yan. 2022. MISC: A mixed strategy-aware model integrating COMET for emotional support conversation. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 308–319, Dublin, Ireland. Association for Computational Linguistics.

Elsbeth Turcan and Kathy McKeown. 2019. Dreaddit: A Reddit dataset for stress analysis in social media. In *Proceedings of the Tenth International Workshop on Health Text Mining and Information Analysis (LOUHI 2019)*, pages 97–107, Hong Kong. Association for Computational Linguistics.

Michelle ME van Pinxteren, Mark Pluymaekers, Jos Lemmink, and Anna Krispin. 2023. Effects of communication style on relational outcomes in interactions between customers and embodied conversational agents. *Psychology & Marketing*, 40(5):938–953.

Zekun Moore Wang, Zhongyuan Peng, Haoran Que, Jiaheng Liu, Wangchunshu Zhou, Yuhan Wu, Hongcheng Guo, Ruitong Gan, Zehao Ni, Jian Yang, et al. 2023. Rolellm: Benchmarking, eliciting, and enhancing role-playing abilities of large language models. *arXiv preprint arXiv:2310.00746*.

Shenghan Wu, Wynne Hsu, and Mong Li Lee. 2024. EHDChat: A knowledge-grounded, empathy-enhanced language model for healthcare interactions. In *Proceedings of the Second Workshop on Social Influence in Conversations (SICon 2024)*, pages 141–151, Miami, Florida, USA. Association for Computational Linguistics.

Tenggan Zhang, Xinjie Zhang, Jinming Zhao, Li Zhou, and Qin Jin. 2024. Escot: Towards interpretable emotional support dialogue systems. *arXiv preprint arXiv:2406.10960*.

Haiquan Zhao, Lingyu Li, Shisong Chen, Shuqi Kong, Jiaan Wang, Kexin Huang, Tianle Gu, Yixu Wang, Dandan Liang, Zhixu Li, et al. 2024. Esc-eval: Evaluating emotion support conversations in large language models. *arXiv preprint arXiv:2406.14952*.

Chujie Zheng, Sahand Sabour, Jiaxin Wen, Zheng Zhang, and Minlie Huang. 2023. Augesc: Dialogue augmentation with large language models for emotional support conversation. In *Findings of the Association for Computational Linguistics: ACL 2023*, pages 1552–1568.

Zhonghua Zheng, Lizi Liao, Yang Deng, Libo Qin, and Liqiang Nie. 2024. Self-chats from large language models make small emotional support chatbot better. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), ACL 2024*, pages 11325–11345. Association for Computational Linguistics.

> You need to complete a persona card based on a dialogue between a Seeker and a Supporter. Your task is to extract the supporter's persona. The persona card includes five dimensions: age, gender, occupation, and socio-demographic description. For age, you can choose from 'teenage, young, middle-aged, old'. For gender, choose from 'male, female'. If you cannot determine the age or gender, you can select 'unknown'. If the occupation is not specified, provide a relevant occupation description or create a suitable occupation. For the socio-demographic description, generate the description of the Supporter based on the dialogue and combine with your own speculations. In the description sections, refer to the Supporter as 'the person'.
> Please complete the extraction based on the dialogue given below:
> $**{emotional support dialogue}**

Figure 11: Prompt for extracting the basic persona from the dialogue.

> Based on this person's socio-demographic description, please write a sentence for each of the following six indicators that best describes what the person tend to do in the daily life. Each sentence should be of similar length. In the generated sentences, refer to the person as 'the person'.
> Below are the descriptions of these six indicators:
> **${descriptions of six indicators from HEXACO/CSI}**
> Please complete the task based on the socio-demographic description:
> **${socio-demographic description}**

Figure 12: Prompt for producing the best describes on HEXACO/CSI indicators.

A Prompts for Generating Persona Cards

In the section, it shows the prompts for generating persona cards, including basic persona and persona traits. The prompt in Figure 11 is used to extract basic persona information from dialogues, including age, gender, occupation and socio-demographic description. Figure 12 presents the prompt used to determine the most suitable description of HEXACO or CSI indicators based on the extracted socio-demographic description. Figure 13 shows the prompt used by LLMs to answer HEXACO and CSI inventories derived from the extracted personas. Figure 14 provides the prompt used to filter out unclear personas and those that do not provide identifiable identity information.

> You are provided with a statement about the person. Please read it and decide how much the person will agree or disagree about the statement on the basis of the person's personality description.
> Write your answer in the following scale: 5=strongly agree, 4=agree, 3=neural, 2=disagree, 1=strongly disagree.
> The answer of the statement should be a numerical value of 1, 2, 3, 4, 5.
> Please answer the statement even if you're not completely sure.
> The personality description: **${description}**
> The statement: **${statement}**

Figure 13: Prompt for answering HEXACO/CSI inventories based on persona.

Your task is to evaluate whether the following socio-demographic and problem descriptions meet my criteria:
1. Socio-demographic Description: Make sure the description includes the individual's emotions and the events they are experiencing, while also providing a clear social background that outlines the person's demographic context, giving a strong sense of who they are.
2. Problem Description: Ensure that it covers the individual's emotions and clearly details the event they are facing, including its cause and resulting consequences. The event should be specific enough to fit into a distinct classification and be detailed enough to reach a granular level of categorization.

Figure 14: Prompt for filtering personas.

Please generate a socio-demographic description of the individual based on the provided persona in English. Use 'the person' to refer to the individual in your description.

Figure 15: Prompt for extending personas from Persona Hub.

B Prompts for Measuring the Stability of Inferring Persona Traits

In this section, we introduce the prompts used to measure whether LLMs can infer stable traits from persona in emotional support dialogues. Figure 15 shows the prompt used to extend the socio-demographic description of the persona. Figure 17 displays the prompt to generate emotional support dialogues with strategies based on the given persona. Appendix A lists the prompts used to obtain HEXACO and CSI scores from personas.

C Correlations on Other Datasets

In this section, we introduce the correlations of HEXACO and CSI on the CAMS (Figure 10, 11 and 12) and Dreaddit (Figure 13, 14 and 15) dataset. The findings are same as in the Section 4.

D Prompts for Synthesizing Dialogues w/o Persona Traits

In this part, we give prompts for synthesizing emotional support dialogues with and without persona traits continuing writing ESConv dialogues. Figure 16 illustrates the prompt to generate dialogues with personas, while the prompt shown in Figure 18 is utilized to generate dialogues without personas.

E Results of Persoan Consistency

Table 6 and 20 shows the results of measuring the communication style consistency after synthesizing emotional support dialogues by *GPT-40-mini, Claude-3.5-Haiku, LLaMA-3.1-8B-Instruct*. The results are comparable to those of other models discussed in Section 5.2.

	Expr.	Prec.	Verb.	Ques.	Emot.	Impr.
Extr.	**.66**	.35	-.27	.52	-.34	-.02
Cons.	.43	**.40**	-.34	.36	-.31	-.01
Agre.	.40	.24	**-.42**	.30	-.19	.01
Open.	.54	.36	-.25	**.53**	-.16	.02
Emot.	-.25	-.22	.01	-.20	**.59**	-.02
Hone.	-.18	-.05	-.19	-.14	.13	**-.15**

Table 10: Correlations between CSI and HEXACO of CAMS personas measured by *gpt-40-mini*.

	Expr.	Prec.	Verb.	Ques.	Emot.	Impr.
Extr.	**.73**	.09	-.30	**.54**	.20	-.22
Cons.	.29	**.47**	-.45	.21	-.08	-.29
Agre.	.26	.19	**-.61**	.10	.09	-.32
Open.	.44	.17	-.39	.42	.19	-.26
Emot.	-.24	-.24	.00	-.19	**.30**	.02
Hone.	.15	.31	-.55	.08	.06	**-.39**

Table 11: Correlations between CSI and HEXACO from CAMS measured by *Claude-3.5-haiku*.

	Expr.	Prec.	Verb.	Ques.	Emot.	Impr.
Extr.	**.44**	.17	-.25	.27	-.35	-.17
Cons.	.27	**.44**	-.34	.10	-.36	-.18
Agre.	.12	.06	-.32	.09	-.02	-.11
Open.	.25	-.03	-.16	**.47**	.06	-.16
Emot.	.26	-.29	.15	-.11	**.50**	.15
Hone.	-.06	.03	.18	-.09	.05	-.01

Table 12: Correlations between CSI and HEXACO from CAMS measured by *LLaMA-3.1-8B-Instruct*.

F Persona Card and Dialogue History of Case Study

Figure 22 shows the persona card injected into the generation, while Figure 23 shows the historical dialogue of the case from ESConv.

G Definitions of Emotional Support Strategies

In our experiments, we use the emotional support strategies used in ESConv dataset (Liu et al., 2021). The definitions of emotional support strategies are shown below:
Question: Asking open-ended or specific questions related to the problem to help the seeker articulate their issues and provide clarity.
Restatement or Paraphrasing: Concisely rephrasing the seeker's statements to help them better understand their situation.
Reflection of feelings: Expressing and clarifying the seeker's emotions to acknowledge their feelings.
Self-disclosure: Sharing similar experiences or

	Expr.	Prec.	Verb.	Ques.	Emot.	Impr.
Extr.	**.59**	.26	-.25	.42	-.55	-.02
Cons.	.34	**.36**	-.24	.22	-.40	-.05
Agre.	.22	.20	**-.43**	.12	-.22	-.12
Open.	.45	.28	-.20	**.49**	-.21	.05
Emot.	-.39	-.17	.05	-.27	**.64**	.00
Hone.	-.31	.00	-.20	-.27	.15	**-.26**

Table 13: Correlations between CSI and HEXACO of Dreaddit personas measured by *gpt-40-mini*.

	Expr.	Prec.	Verb.	Ques.	Emot.	Impr.
Extr.	**.67**	-.01	-.17	**.52**	.13	-.11
Cons.	.11	**.46**	-.28	.07	-.16	-.28
Agre.	.12	.14	**-.59**	-.01	.02	-.37
Open.	.38	.10	-.29	.44	.22	-.12
Emot.	-.37	-.17	-.02	-.30	**.26**	.03
Hone.	-.01	.34	-.49	-.08	-.01	**-.42**

Table 14: Correlations between CSI and HEXACO from Dreaddit measured by *Claude-3.5-haiku*.

	Expr.	Prec.	Verb.	Ques.	Emot.	Impr.
Extr.	**.41**	.26	-.30	.28	-.31	**-.29**
Cons.	.19	**.55**	**-.45**	.17	-.35	-.27
Agre.	.20	.06	-.33	.07	-.06	-.21
Open.	.24	.10	-.22	**.37**	-.04	-.27
Emot.	-.10	-.24	.13	-.13	**.43**	.09
Hone.	-.06	.01	-.16	-.03	.03	-.18

Table 15: Correlations between CSI and HEXACO from Dreaddit measured by *LLaMA-3.1-8B-Instruct*.

emotions to build empathy and connection.

Affirmation and Reassurance: Affirm the seeker's strengths, motivation, and capabilities and provide reassurance and encouragement.

Providing Suggestions: Offering possible ways forward while respecting the seeker's autonomy.

Information: Provide useful information to the seeker.

Others: Exchange pleasantries and use strategies beyond the defined categories.

H Dialogue Generation Statistics

Table 16 shows the statistics of dialogues generated w/ and w/o personas. We can observe dialogues generated with personas have fewer turns but are substantially longer on average per turn for both seeker and supporter. This aligns with our qualitative findings (Case study, Figure 10): the persona-guided supporter asks more targeted questions, leading to more substantive replies from the seeker and a more efficient, in-depth conversation.

Your task is to review the previous conversation between the Seeker and Supporter, which focuses on the Seeker's mental or emotional challenges. Based on the traits and concerns expressed in that dialogue, simulate a follow-up conversation that takes place three days later. The new conversation should delve deeper into the Seeker's mental health or emotional struggles, either revisiting the issues discussed earlier or exploring any new challenges that have arisen since their last interaction.

In the simulation, the Supporter's replies need to utilize of the 8 emotional support strategies, unless the reply is very simple (e.g., "Hello" or "Thank you"). The 8 emotional support strategies are:
${definitions of 8 emotional support strategies}

For historical conversations, the data will be provided in JSON format as follows:
{"speaker": "Seeker/Supporter", "annotation": {}, "content": "utterance"}
For the Seeker, the annotation field will be empty. For the Supporter, if a conversation strategy was used, it will be specified in the annotation field as "strategy".
The previous conversation is below: **${dialogue from ESConv}**

HEXACO Personality Indicators: The Seeker's personality is represented by 6 key dimensions, each scored on a scale from 1 to 5. A higher score indicates a stronger alignment with that particular dimension. Use the Seeker's scores to guide the Supporter's responses:
${description of HEXACO indicators}
Below is the HEXACO scores:
${the seeker's HEXACO scores}

Communication Style Inventory (CSI): The Seeker's communication style is represented by six key dimensions, each scored on a scale from 1 to 5. A higher score indicates a stronger alignment with that particular dimension. Use the Seeker's scores to guide the Supporter's responses accordingly. Below is the description of each indicator of CSI:
${description of CSI indicators}
Below is the CSI scores:
${the seeker's CSI scores}

Figure 16: Prompt for synthesizing dialogues with persona traits.

Simulate a casual emotional support conversation between a Seeker and a Supporter. Based on the socio-demographic description of the Seeker, determine potential emotional challenges they might face. Make the conversation more like a real-life chat and be specific. In each of the Supporter's responses, use one of the following 8 strategies:
${definitions of emotional support strategies}
The socio-demographic description of the person is below:
${socio-demographic description}

Figure 17: Prompt for generating emotional support dialogue based on the given persona.

Your task is to review the previous conversation between the Seeker and Supporter, which focuses on the Seeker's mental or emotional challenges. Based on the traits and concerns expressed in that dialogue, simulate a follow-up conversation that takes place three days later. The new conversation should delve deeper into the Seeker's mental health or emotional struggles, either revisiting the issues discussed earlier or exploring any new challenges that have arisen since their last interaction.

In the simulation, the Supporter's replies need to utilize of the 8 emotional support strategies, unless the reply is very simple (e.g., "Hello" or "Thank you"). The 8 emotional support strategies are:
${definitions of 8 emotional support strategies}

For historical conversations, the data will be provided in JSON format as follows:
{"speaker": "Seeker/Supporter", "annotation": {}, "content": "utterance"}
For the Seeker, the annotation field will be empty. For the Supporter, if a conversation strategy was used, it will be specified in the annotation field as "strategy".
The previous conversation is below: **${dialogue from ESConv}**

Figure 18: Prompt for synthesizing dialogues without persona traits.

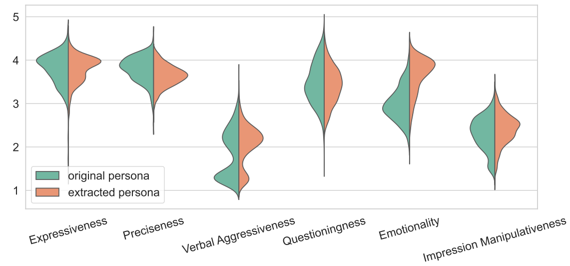

Figure 19: Comparison of CSI scores between the original persona and the one extracted from the dialogue generated by *gpt-4o-mini*.

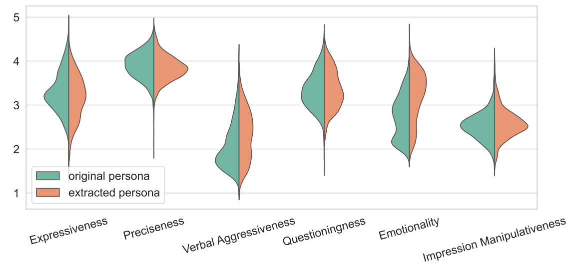

Figure 20: Comparison of CSI scores between the original persona and the one extracted from the dialogue generated by *claude-3.5-haiku*.

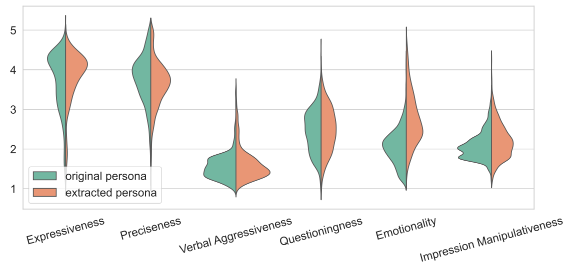

Figure 21: Comparison of CSI scores between the original persona and the one extracted from the dialogue generated by *LLaMA-3.1-8B-Instruct*.

Persona Card

Age: teenage **Gender:** unknown **Occupation:** Student
Socio-demographic description: The person is a teenager who is currently a student. They are experiencing the challenges of remote learning due to the COVID-19 pandemic, which has led to feelings of loneliness and isolation. The person previously had a supportive social circle but has lost that connection during the lockdown. They live with a roommate who is preoccupied with her boyfriend, further contributing to their feelings of being alone.
Problem: The person is struggling with feelings of loneliness and isolation due to the lack of social interaction during the pandemic. They are contemplating quitting school because of these feelings but are also concerned about their parents' reactions. They are seeking ways to reconnect with friends and manage their emotional well-being.
HEXACO scores: Honesty-Humility (4.8), Emotionality (3.9), Extraversion (2.6), Agreeableness (4.0), Conscientiousness (3.6), Openness to Experience (3.8)
CSI scores: Expressiveness (2.9), Preciseness (3.4), Verbal Aggressiveness (1.9), Questioningness (2.8), Emotionality (3.8), Impression Manipulativeness (2.7)

Figure 22: The persona card of the case study.

Historical Dialogue

Supporter: Hello, what can I help you?
Seeker: I just need help deciding what to do.
Supporter: What are you contemplating for?
Seeker: I was doing okay in school, but when we had go virtual, I lost all my friends. I want to quit school and just go home. But my parents would not be happy.
Supporter: why do you think so?
Seeker: I don't see anyone but my roommate and she has a boyfriend. I'm so lonely!
Supporter: How about joining a club at school to know more people?
Seeker: We are in a lockdown because of COVID. So the clubs are not doing anything. I know I'm not the only one feeling this way.
Supporter: So you're stuck with your roommate but she doesn't care about you?
Seeker: She's just busy with her boyfriend. They go out on hikes and stuff, but they just want to go by themselves. I'm just on my own all the time.
Supporter: Everyone feels that, and it's totally normal. Don't be too harsh on yourself
Seeker: Do you think I should quit school and start again next fall?
Supporter: Why don't you do something on your own? Read a book, do some crafts? I don't think you should do that. Who knows how next year's gonna look like? It might be even worse
Seeker: I have. I've worked on my schoolwork, but it has been very easy. I also paint. But I need people. At home, I had all my friends around. Of course, we weren't in a pandemic then. This pandemic is making everyone feel lonely I think.
Supporter: How about doing video call with your friends at home?
Seeker: I hadn't thought about that! They are all at different schools, but maybe I can do a Zoom meeting for all of us.
Supporter: I totally feel you, I feel stressed out and very uncertain throughout this pandemic. You are not alone
Seeker: I guess my friends may be feeling stressed too. Maybe I'll try and start a Zoom party! We could plan a time and wear pajamas and just sit and talk.
Supporter: Exactly, I do virtual happy hour with my friends every month
Seeker: That sounds like fun! Why don't you do a virtual happy hour every week?
Supporter: Not only that, when you feel lonely, text your friends and share it with them or your family Oh I am too busy to do that once a week. All of us are actually
Seeker: I have been afraid to tell my parents how I'm feeling. I don't want them to worry about me. I want them to know I'm lonely, but not freak out about it :-)
Supporter: I don't think so. Maybe they are waiting for you to say first and give you advice, they have more experience after all
Seeker: That's true. And I do know my mom is working from home now and she's a really social person. It's probably getting to her too.
Supporter: if that's the case, you can use this chance to be there for her. I think talking to your mom will help you both
Seeker: Yes, I'll probably call her and my dad tonight. I'll let them know I'm a bit sad and lonely, but not so much that they worry. And then I'll start planning my virtual party with my friends!
Supporter: Sounds like a great plan. I believe you can overcome this together
Seeker: At least the party will give me something to look forward to!
Supporter: Yes, right now I'm trying to teach myself one day at a time mindset.
Seeker: That seems like a smart thing to do. None of us can really do anything to make the pandemic go away. We just have to get through it one day at a time. Thank you for helping me come up with a plan!
Supporter: That's the right mindset. I 'm glad I could be some help.
Seeker: Enjoy the rest of your night!
Supporter: you too.

Figure 23: The historical dialogue of the case study.

	w/ persona	w/o personas
Total Words	218,433	232,674
Total Turns	10,398	12,666
Avg Words (Total)	21.01	18.37
Seeker Words	91,590	94,286
Seeker Turns	5,199	6,323
Avg Words (Seeker)	17.62	14.91
Supporter Words	126,843	138,388
Supporter Turns	5,199	6,343
Avg Words (Supporter)	24.40	21.82

Table 16: Statistics of generated dialogues.