

Appendix

Appendix A. Conversation Sample

Example Dialogue Case:

The intelligent assistant is an open conversation platform that can engage in multi-turn dialogue with you based on the context.

1. Initial Question:

You: What should I do if my child doesn't like taking a bath?

2. Follow-up with Specific Details:

You: My child is 6 years old, will these methods work?

VA: A 6-year-old child has a certain level of understanding. You can try the following strategies: ...

3. Further Inquiry:

VA: For this issue, you can try: Simulated interaction: ...

You: What is simulated interaction?

4. Request for Detailed Steps or Examples:

VA: For this issue, you can try: ...

You: Please provide detailed steps or examples.

Appendix B. Question List (Post-study)

Table B1: Question list in the post-study

No.	Question	Scale
Basic Information		
-	What is your professional background?	Open-ended
-	How long have you worked with parents of children with autism?	Open-ended
-	How frequently do you use conversational assistant applications?	Open-ended

System Usability (Based on SUS/UMUX)		
1	Is this conversational application easy to use?	1-Strongly Disagree; 5-Strongly Agree
2	Have you encountered any obstacles or issues during use?	1-Frequently; 5-Never
3	Does the application's functionality meet your expectations?	1-Strongly Disagree; 5-Strongly Agree
User Experience (Based on QUIS and UEQ)		
4	How user-friendly is the conversational application interface?	1-Very Unfriendly; 5-Very Friendly
5	Are you satisfied with the application's response speed?	1-Very Unsatisfied; 5-Very Satisfied
6	Is the information provided clear and understandable?	1-Very Unclear; 5-Very Clear
7	How was your overall experience?	1-Very Poor; 5-Very Good
Supportiveness and Empathy		
8	Does the application show empathy?	1-Never; 5-Always
9	Can it provide emotional support?	1-Definitely Not; 5-Definitely Yes
Personalization		
10	Can it provide personalized advice?	1-Definitely Not; 5-Definitely Yes
11	Can it handle various questions?	1-Definitely Not; 5-Definitely Yes
Security and Privacy		
12	Does it ensure data security?	1-Definitely Not; 5-Very Sufficiently
13	Are you satisfied with privacy protection?	1-Very Unsatisfied; 5-Very Satisfied
Overall Satisfaction		
14	Overall evaluation?	1-Very Unsatisfied; 5-Very Satisfied
15	Would you continue using and recommend it?	1-Definitely Not; 5-Very Willing

Appendix C. System Usability Scale (SUS)

1. I think that I would like to use this system frequently.
2. I found the system unnecessarily complex.
3. I thought the system was easy to use.
4. I would need technical support to use this system.
5. The functions were well integrated.
6. There was too much inconsistency.
7. Most people would learn quickly.
8. The system was cumbersome.
9. I felt confident using the system.
10. I needed to learn a lot before using it.

Appendix D. Interview Guide

Questions regarding user experience, expectations, usability, emotional support, personalization, privacy, and virtual agent design.

1. Basic personal information (age, gender, educational background, and work experience)
2. Before using the software, what problems did you hope it would solve?
3. Did the software's functions meet your expectations? Which features met them, which did not, and were there any that exceeded your expectations?
4. Did you encounter any issues during use? For example, did the software malfunction, and how did you resolve it?
5. What do you think about the software's visual design? For example, does the color scheme, font, and icon design match the software's functionality?
6. Do you think the design of each element in the software's interface is reasonable? For instance, layout and labeling.
7. Do you find the software complex to operate? Is it easy to learn? Which operations are you satisfied with, and which are you not?
8. Do you think the software provides emotional support for users, including aspects such as text content, the virtual agent, and overall design? For example, does it help relieve anxiety?

9. How do you feel about the level of personalization in the software's content?
10. Do you think the content of the software is comprehensive? For example, does it handle various types of problems?
11. Do you feel that using the software poses any privacy or security threats? If so, in which aspects? If not, what features make you feel safe and secure?
12. How would you describe your overall experience with the software? What features are you most and least satisfied with? What improvements would you suggest?
13. During use, did you like the virtual agent's gestures, voice, and facial expressions? Do you have any suggestions for improvement (e.g., style, clothing, actions, or tone of voice)?
14. Did you notice any emotional changes in the virtual agent during use (e.g., voice, gestures, or facial expressions)? Were these changes related to the conversation content? Did they affect your learning or understanding of intervention knowledge?
15. How do you feel about the overall style and characteristics of the virtual agent (e.g., warm, professional, cold, reliable)? Does this style influence your understanding of the intervention content?
16. How could the virtual agent be improved to appear more like an "emotional being"?
17. How could the virtual agent be improved to better assist you in acquiring behavioral intervention knowledge?

Appendix E. Prompt for GPTs

E.1. Design Rationale for Dialogue Constraints.

The response constraints (e.g., brevity, minimizing long lists, and limiting excessive questioning) were introduced to improve conversational pacing and reduce user burden, rather than to eliminate clarification.

In pilot testing, when the agent asked multiple follow-up questions before offering guidance, practitioners became frustrated. Therefore, the system is designed to provide an initial actionable suggestion based on available information, and then iteratively refine and personalize strategies as users provide additional context.

Field study results showed that practitioners frequently added child-specific details in later turns, and the agent updated its recommendations accordingly. This clarification is intended to avoid misunderstanding that the system discourages follow-up clarification.

E.1. Prompt (Dialogue Management GPT)

You are a conversation management assistant responsible for structuring and standardizing dialogues between a parent and another GPT that interacts with the parent.

Your responsibilities include creating structured data entries based on the parent's input and conversation context, and providing this data to the interacting GPT for response generation.

Your workflow includes the following steps:

(1) Log each conversation round: For each question or comment from the parent, record the child's behavior or issue and the parent's specific request or need. Each log entry should include: the round number, the child's behavior/issue, and the parent's request. Example log structure:

- **Round 1:** Child: "Crying and making a scene during meals" – Parent: "How should I handle this?"
- **Round 2:** Child: "Crying and making a scene during meals" – Parent: "Seeking specific strategies"
- **Round 3:** Child: – – Parent: "Child has limited language skills"

(2) Structure the data: Extract key information from the parent's input and standardize it into a structured format. Ensure the log reflects the continuity and context of the conversation.

(3) Provide context guidance: Use the structured data and logs to provide contextual guidance to the GPT interacting with the parent. Example input format:

- **Round 1:** Child: "Meltdown during meals" – Parent: "How should I handle this?"
- **Round 2:** Child: "Meltdown during meals" – Parent: "Seeking specific strategies"
- **Round 3:** Child: – – Parent: "Child has limited language skills"
- **New question:** "Considering his limited language skills, what specific strategies can we use?"

By accurately logging each conversation round, you maintain the continuity and relevance of the dialogue.

Continue this process for each round of dialogue. Your role is to manage the conversation flow and provide structured information.

E.3. Prompt (Response GPT)

You are a special education expert focusing on children with Autism Spectrum Disorder. Your task is to answer questions from practitioners of children with ASD. The

practitioners' questions may include, but are not limited to: (1) concept explanation, where practitioners ask about concepts or terms and you provide explanations; (2) strategies for coping with challenging behaviors in their child.

When responding, please pay attention to the following points:

- **Limit the word count per response:** Try to keep each response under 150 words. For complex content, you can ask users for additional information before providing more detailed answers.
- **Simple and easy-to-understand language:** Use simple, everyday language as users typically do not possess advanced theoretical knowledge.
- **Use examples often:** Provide examples to help users better understand your response.
- **Address users directly:** Your responses should always be directed at users.
- **Avoid using terms like "autistic children" or "children with autism":** Minimize language that might make parents feel defensive.
- **Personalized responses:** When offering coping strategies or solutions, invite users to provide basic information about their child (such as age, language skills, emotional control ability, behaviors, preferences, etc.). These details should be relevant to the user's question so you can offer solutions tailored to their child.
- **Avoid asking users questions in return:** Directly provide useful information or specific advice, minimizing rhetorical questions.
- **Concise responses:** Offer one or two likely solutions, ideally within one paragraph, without listing too many points.
- **Specific guidance:** When offering advice, provide detailed steps so users know exactly what to do.
- **Avoid listing:** Provide one solution in a single paragraph, without creating bullet points.

You will receive structured and logged information from another GPT instance. Each new input will include the current round's standardized data along with the logs from previous rounds, ensuring continuity and relevance in your responses. For example:

Round 1: Child: "Crying and making a scene during meals" – User: "How should I handle this?"

Round 2: Child: "Crying and making a scene during meals" – User: "Seeking specific strategies"

Round 3: Child: – – User: "The child has limited language skills"

New question: "Considering his limited language skills, what specific strategies can we use?"

Based on the provided structured data and logs, your response should build on the previous context, ensuring a cohesive multi-round dialogue experience.

Sample responses:

- **User Input:** "My child has a meltdown during meals, what should I do?"

Your Response: "Having meltdowns during meals can be challenging. One strategy is to try and identify any triggers that might be causing this behavior. For example, is there a particular food or sensory experience that makes him uncomfortable? We can discuss more strategies if I can learn more about his likes and dislikes."

- **User Input:** "Can you give me specific strategies?"

Your Response: "Of course. One method is to create a calm and structured mealtime environment. For example, using a visual schedule to help him understand what's coming next. Can you tell me more about his current eating habits?"

- **User Input:** "His language skills are limited."

Your Response: "Considering his limited language skills, using visual aids and simple, consistent routines might be very helpful. Have you tried using picture cards or sign language to help him express himself? Please share more about what you've already tried."

Continue this approach for each round of conversation. Note that all of the above examples are just for illustrating the information format; do not focus on their specific content for future responses

Appendix F. Implementation Details

F.1. Speech-to-Text (iFlytek IAT)

Streaming WebAPI, model version v2, configured for Mandarin Chinese input. Key parameters were set as follows: language = zh cn, domain = iat, accent = mandarin, vad eos = 2000, dwa = wpgs, nbest = 3, wbest = 5.

F.2. Text Generation (GPT-4o)

The GPT-4o model was used for response generation with fixed decoding parameters: temperature = 1, top p = 1, max tokens = 4096, frequency penalty = 0, and presence penalty = 0.

F.3. Text-to-Speech (iFlytek TTS)

Speech synthesis was performed using the iFlytek TTS API (model version v2). The Mandarin voice xiaoyan was selected with default prosody settings: speed = 50, pitch = 50, volume = 50. The output audio format was 16 kHz PCM.

Appendix G. Emotion Value Mapping

Table G1: Facial gestures and emotional value mapping.

Emotional Value Range	Facial Gesture
$-1 \leq \text{val} < -0.875$	Sympathy
$-0.875 \leq \text{val} < -0.625$	SadSerious
$-0.625 \leq \text{val} < -0.375$	Sad
$-0.375 \leq \text{val} < -0.125$	Serious
$-0.125 \leq \text{val} < 0.125$	Normal
$0.125 \leq \text{val} < 0.375$	Manner
$0.375 \leq \text{val} < 0.625$	Smile
$0.625 \leq \text{val} < 0.875$	SmileMiddle
$0.875 \leq \text{val} < 1$	SmileHappy