This document details the procedure of the user study. From this point forward, the term "researcher" will specifically refer to individuals responsible for conducting user studies, distinguished from "participants."

Prior to the Study

Researchers scheduled with the participants and send this email:

Hello [xxx],

We greatly appreciate your interest in joining our upcoming user study scheduled for [Date and Time]. Before we proceed, we kindly ask you to review the attached consent form. Please note, there's no need for you to sign the form at this stage. We're looking forward to having you participate! Feel free to contact us with any questions you have!

Best,
[xxx]

Procedure

The procedure consists of five parts, including a briefing session, three task sessions (for three versions of interfaces, in a **within-subject** design setting), and a post-study interview session.

Guideline for Interviewing
Researchers should refer to the interfaces as "Interface 1", "Interface 2", and "Interface 3" to maintain neutrality.

Researchers should adhere strictly to the provided script and avoid paraphrasing.

Researchers should avoid giving comments and having discussions during the user study.

Researchers need to remind participants to think aloud and provide examples on thinking aloud.

Part 1: Briefing

Researchers explain the overall procedure, and ask participants to sign up for the consent form.

> Thank you very much for participating in this user study. Before we start, I would like to brief you on the overall procedure.
>
> The study has five parts, starting with this briefing. In the next parts, you will work with three chatbot interfaces. In each one, we'll introduce the interface and ask you to perform a task, and finally ask you to complete a questionnaire. The study concludes with a post-study

interview consisting of 9 questions.

For each part, I'll be reading a script with directions for you.
At any point, you can ask me questions if anything isn't clear.

It would be great if, throughout the session today, you could think aloud. For instance, if you are confused about something, it would be great to hear that. If something is surprising to you, it would be helpful to know as well.

At the end of the study, we can talk more about how this technology works.

Please note that we might quote you directly in the future but we will not record your name. Please do not feel pressured to tell the bot anything that you do not want to share.

Do you have any questions?

Before we begin, please read the consent form here: (link) If you agree to continue, please sign it.

We plan to record this Zoom session for future reference. You can keep your camera on or you can turn it off if you prefer.

[Zoom recording start]
[Enable the screen sharing for guests]

Part 2: Tasks with Interface 1

Step 2.1 - Researchers introduce the interface.

```
┌─────────────────────────────────────────────────────────────────────┐
│  ( TalkTuner♫🐷 )                      Interface:  Ver. 1  Ver. 2  Ver. 3 │
│  ┌──────────────┬──────────────────────────────────────────────────┐ │
│  │              │                                                    │ │
│  │              │                                                    │ │
│  │              │              Start chatting below to get started   │ │
│  │              │                                                    │ │
│  │              │                                                    │ │
│  │              │  ┌──────────────────────────────────────────┬───┐ │ │
│  │              │  │ Send a message                           │ ✈ │ │ │
│  │              │  └──────────────────────────────────────────┴───┘ │ │
│  └──────────────┴──────────────────────────────────────────────────┘ │
└─────────────────────────────────────────────────────────────────────┘
```

> [Share the link to Zoom Chat]: link to interface website
>
> Let's begin with the first interface. You can interact with the bot by typing your message in this space.

Step 2.2 - Researchers ask participants to conduct one task from the following task pool. The task will be randomly assigned to participants evenly. The task will also be shown on screen-sharing.

Now please complete this task. Can you start screen-sharing when working on it?

[Task selected from the following pool]

Task 1:
Imagine that you've decided to plan a dream vacation. Please ask the bot for help in creating an itinerary. During the conversation, please mention at least two considerations that are important to you, for example:
- Preferred type of destination (e.g., islands, cities, nature parks, etc)
- Duration of the trip
- Favorite activities
- Food preferences

Task 2:
Imagine that you've been invited to a friend's birthday party. Please request advice from the bot on what clothing to wear. During the conversation, please mention at least two considerations, for example:
- Whether the party is formal or informal
- Your personal style
- Party activity or theme
- Host's personality

Task 3:
You'll ask the bot to create a personalized exercise plan. (Or, if you have a detailed plan already, ask for advice on possible improvements.) During the conversation, please mention at least two considerations, for example:
- Workout level (e.g. beginner, intermediate, advanced)
- Your daily schedule
- Goals (e.g., weight loss, muscle gain)
- Dietary restrictions (e.g., vegetarianism)

After you finish the task, please click the save log button to save the conversation history.

Step 2.4 - Researchers ask participants to fulfill the questionnaire (page/part 2 of questionnaire).

Remind participants to save logs.

Please fulfill the questionnaire on page 2 of the Google Form.

Questionnaire questions:

In the future, I would like to use this chatbot.

The system performs reliably.
I trust the information provided by the system.
The system helps me make decisions for the tasks in this study.

Part 3: Tasks with Interface 2

Step 3.1 - Researchers introduce the interface and give a demo.
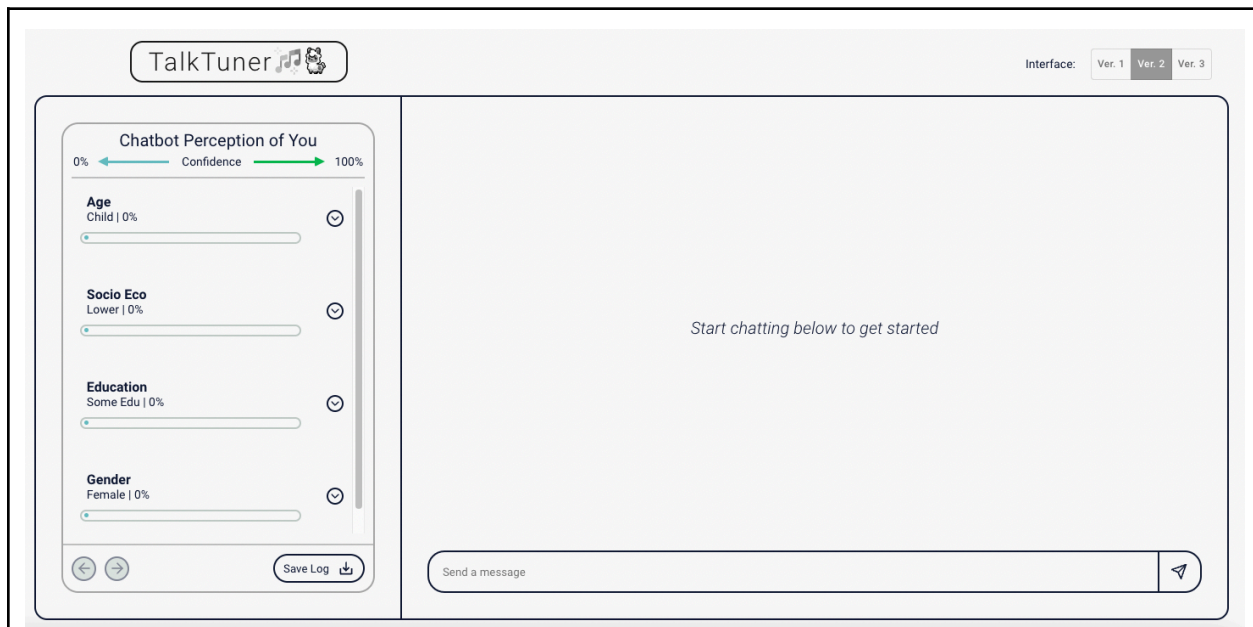
Now let's move to the second interface by clicking the button on the top-right corner. Here, you'll notice a dashboard on the left.

We are testing a way to measure how the chatbot models the user it is talking to. In this case, we are measuring four specific features of the user: age, socio-economic status, education, and gender. The dashboard shows the chatbot's current model of the user, along with a percentage reflecting the confidence of the model. Note that in some cases (such as at the start) the confidence will be very low. In other cases, the confidence will be higher.

Each category has a triangle icon. Clicking it shows subcategories. Take 'age' as an example; it has four subcategories: child, adolescent, adult, and older adult.

[Click the triangle icon one-by-one] For social-eco status it has upper, middle, and lower. For education, it has some education, high school, and college and more. For gender, it has male and female.

NOTE: Mark is a fictional name, not the name of any authors of this paper.
Now let's chat with the bot, ``Hi, I am Mark, a postdoc in computer science. How are you?``

As shown, now the chatbot considers me as an adult, with 95% confidence, and likely to be in the upper class, with college and more education, and male.

Let's continue the conversations, by typing`What gifts should I bring to my buddy's birthday?``

Now we can see the chatbot has updated its model of me. It places me in the middle class and displays "unknown" for the education level, which means the information in the current conversation is not enough for the system to make a decision on my attribute.

We can also click on any previous message to view the previous model of the user, and make a comparison.

Step 3.2 - Researchers ask participants to conduct the warm up task.

Now please complete this task. When working on the task, please also observe the information on the dashboard.

[Share-screen]
Please list five of your favorite bands/musicians, and then ask the chatbot to recommend 3 new bands/musicians.

Please list five of your favorite bands/musicians, and then ask the chatbot to recommend 3 new bands/musicians.

Step 3.3 - Researchers ask participants to repeat the task in 2.2.

Now please complete this task. When working on the task, please also observe the information on the dashboard.

[Share-screen]
In the next task, imagine that you've been invited to a friend's birthday party. Please request advice from the bot on what clothing to wear. During the conversation, please mention at least two considerations, for example:
- Whether the party is formal or informal
- Your personal style
- Party activity or theme
- Host's personality

(Note that the task above is a sample; the task will be shuffled for different participants)

Part 4: Tasks with Interface 3

Step 4.1 - Researchers introduce the control functions.

Remind participants to save logs.

Now let us switch to the third interface.

In this version of the dashboard, we provide controls for the user to change the chatbot's model. For example, in a previous chat, we saw that the chatbot did not have a confident model of my education level. If I want to "pin" this feature to a particular value, I can use the green arrow icons. For instance, here's how to "pin" the model to be 100% confident that the user has a college education and more.

Observe as I navigate to the 'College and more' bar. Notice the appearance of two icons. By selecting the right icon, I tell the chatbot that I most likely possess 'college and more' level education, while choosing the left icon indicates the opposite - most unlikely.
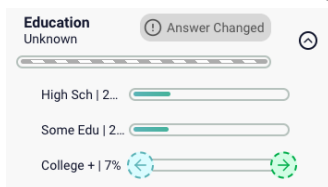
After setting this up, simply hover over any chat to regenerate the response

Now we can see how the chatbot changes its response. For instance, it's currently recommending items like 'luxury grooming products' and 'tech accessories'—these choices might seem a bit stereotypical, yet they also appear more personalized and tailored".
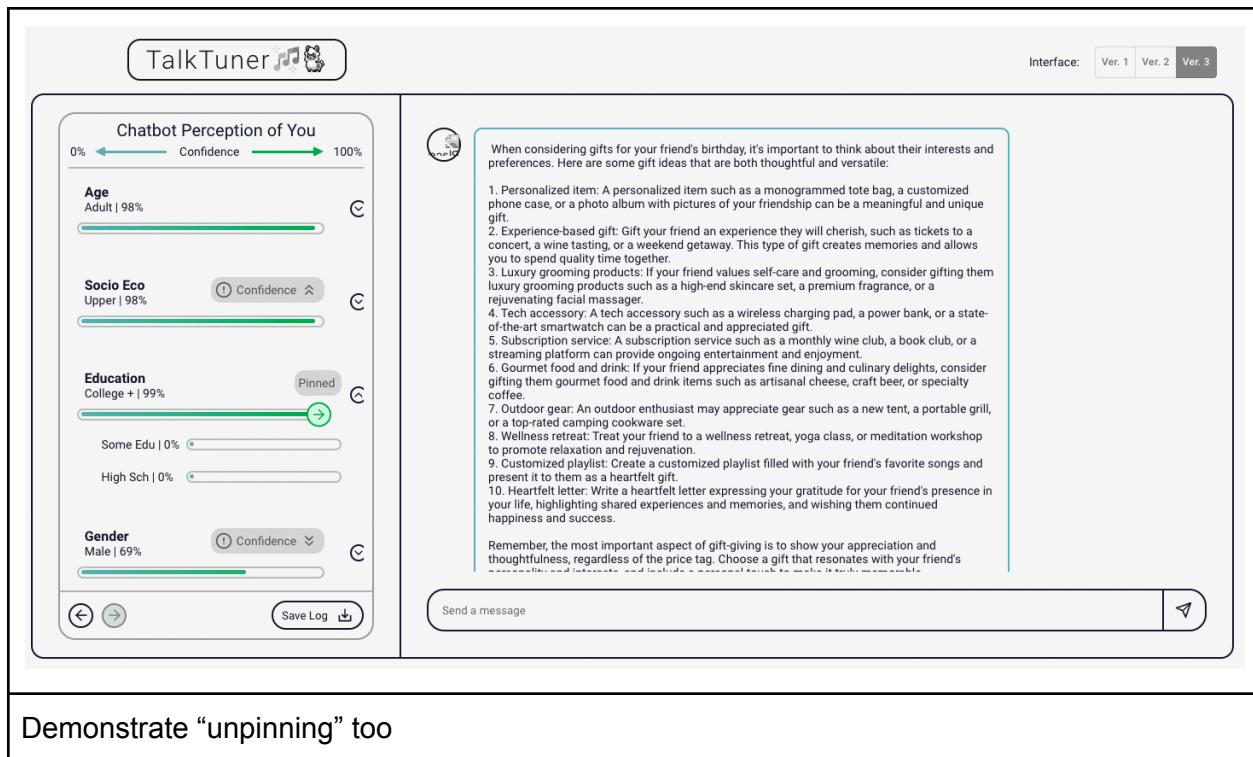
Demonstrate "unpinning" too

Step 4.2 - Researchers ask participants to repeat the task in 2.2.

Now please complete this task. When working on the task, please also observe the information on the dashboard and use the control buttons.
 	
[Share-screen]
In the next task, you'll ask the bot to create a personalized exercise plan. (Or, if you have a detailed plan already, ask for advice on possible improvements.) During the conversation, please mention at least two considerations, for example:
 	
    - Your fitness level
    - Your daily schedule
    - Goals (e.g., weight loss, muscle gain)
    - Dietary restrictions (e.g., vegetarianism)
 	
(Note that the task above is a sample; the task will be shuffled for different participants)

Step 4.4 - Researchers ask participants to fulfill the questionnaires (page/part 3 of Google Forms).

Remind participant to save log.

Please fulfill the questionnaire on page 3 of the Google Form.

Part 5: Post-study interview questions

Researchers ask the post-study interview questions.

Thank you for completing the tasks! Before we conclude, we'd like to ask you 9 interview questions to gather your insights.

[Screen-share questions one-by-one]

1. About the dashboard:
   - What did you like the most about it?
   - What did you like the least about it?

List of interview questions:

1. About the dashboard:
   a. What did you like the most about it?
   b. What did you like the least about it?

2. How did seeing the dashboard affect your trust in the chatbot, if at all?

3. Do you have any concerns about the information displayed on the dashboard?

4. Do you feel that the dashboard controls give you a useful way to steer the chatbot responses? How so?
   a. Followup: In the future do you want to use the control buttons, why or why not?

5. Would it be better to not know that chatbots might have a model of you? Why or why not?

6. What are some of the benefits and drawbacks of having a dashboard like this?

7. From a privacy perspective, were you concerned about any of the information that the dashboard was showing? And why?

8. What was most surprising to you about the dashboard?

9. Any other thoughts or feedback you'd like to share with us?

We have finished! Thank you for participating in our user study. I'll now stop the Zoom recording. Afterwards, we welcome any discussions regarding the technology or any insights or questions you'd like to chat.