Combining Closed-Based and Chatbot-Based Questions in Surveys
An Experiment with GPT-3

Pierre Petronin
pierrepetronin@google.com

Mario Callegaro
callegaro@google.com
Conversational surveys help reduce satisficing, but…

“We found that the participants in the chatbot survey, as compared to those in the web survey, were more likely to produce differentiated responses and were less likely to satisfice”

Chatbot surveys are an effective tool to reduce satisficing behavior [Kim, Lee, Gwon, 2019].

“Our collaborator, the market research firm [...] want to explore a hybrid use of a chatbot with a typical online survey where they can use a platform like Qualtrics to ask complex quantitative questions (e.g., Matrix rating) while leveraging a Juji-like chatbot for open-ended questions. “

Example experience

Based on your last experience, how easy or difficult was it to use Google Maps on your phone?

- Very easy
- Somewhat easy
- Neither easy nor difficult
- Somewhat difficult
- Very difficult
In this question, you can talk with a chatbot about your Google Maps experience.

What did you find hard to use or frustrating when using Google Maps?

Type your response here
In this question, you can talk with a chatbot about your Google Maps experience.

What did you find hard to use or frustrating when using Google Maps?

it did not get my location right
In this question, you can talk with a chatbot about your Google Maps experience.

What did you find hard to use or frustrating when using Google Maps?

It did not get my location right

Why didn't Google Maps get your location right?

Type your response here
In this question, you can talk with a chatbot about your Google Maps experience.

What did you find hard to use or frustrating when using Google Maps?
- It did not get my location right

Why didn't Google Maps get your location right?
- It thought I was 200m away from my real location

Were you indoors or outdoors when you were trying to get directions?
- I was just exiting the subway station

I've got all the info I need. Thank you for this conversation!
How can we build a hybrid solution?

Our solution is low on code, can be used in parallel by different researchers.

- **OpenAI** produces follow-up question + toxicity score based on a (question, response) + model reference.
- **node.js** exchanges completion requests with OpenAI, stores API key, ends conversation after 3 exchanges or when toxic content is detected.
- **Qualtrics** embeds node.js app as iframe, sends the initial question to node.js, stores respondents’ answers.

**Training data** = 100 exchanges

330 lines of code

70 lines of code
Survey respondents were randomized into 2 groups to test the effect of a hybrid vs traditional web survey experience. 

1,200 complete responses, desktop respondents, US Panel provider: Cint

Closed-ended questions on recent Google Maps experience

Open-ended question

Chatbot

Chatbot perception questions

1,200 complete responses, desktop respondents, US

Panel provider: Cint
Respondents using the chatbot produced longer responses in total

Open-end: 25 char
Chatbot: 54 char
Getting “one-level deeper”

How do researchers benefit?

Connecting topics

Clarifying ambiguous statements

What did you find hard to use or frustrating when using Google Maps?
> i am old and it takes me longer to learn something and then to remember how to use it, but i am getting better at apps and maps.

What specific aspects of Google Maps could be easier for you to use or use more easily?
> using the little numbers and letter are so hard for me, i wish they were bigger on the phone so it would be easier for me to use

What would make the numbers and letter bigger on your phone?
> larger buttons
How do respondents feel about the experience?

Respondents need to feel understood to become engaged

39% preferred the chatbot experience | 31% had not preference | 30% preferred a traditional open-ended question

Shows “understanding”

“All the responses made sense, it was actually an incredibly good chatbot. It felt very human.”

“it is fast and was specific about what i had just answered”

Engaging

“it makes surveys a little bit more entertaining and gets you out of the mode of surveying and more into the mode of having a conversation overall.”

Repetitive questions / words

“I don't appreciate questions being repeated if the response is not sufficient”

“it just repeated key words”

When does it stop?

“A bit confusing. I didn't know when the interaction was slated to end”
What we’ve learned

We can train a chatbot to follow-up on open-ended responses, and integrate it in web-based surveys. Our solution is small, modular, re-usable.

Using a chatbot resulted in higher quality and more detailed responses. Total response length was ~2x. Increase in specificity may make it easier to code answers.

Experience was overall slightly better for respondents compared to a traditional open-ended. Respondents felt understood, and felt like they were having a “real” conversation.
Future work

**Improve the chatbot**
/ Better probes
/ Allow using different Large Language models

**Better understand impact on responses**
/ Do certain respondents better react to it?
/ Does this reduce satisficing, or extract more info from non-satisficers?

**How might we automate the analysis of conversations?**
/ Topic clustering for conversations?
Try it!


pierrepetronin@