

CS6750 HCI Summer 2021: Project Redesign Chatbot Interface of Agoda

Nan Xiao
nanx@gatech.edu

Abstract—In this paper, we will redesign the chatbot interface of Agoda. Firstly, we will introduce how to get to the Agoda chatbot interface. Secondly, we will plan and execute the initial needfinding to understand the weakness of the existing interface and what features the interface is expected to have. Thirdly, we will perform a heuristic evaluation on the existing interface to understand pros and cons and the reasons. Fourthly, we will conduct the redesign of the interface using the wireframe prototype. Fifthly, we will justify the redesign and explain how the redesigned interface addresses the issues raised in initial needfinding. Finally, we will discuss the evaluation plan of the redesigned interface.

1 INTRODUCTION

Customer service is very costly in travel domain, so most of the OTA (Online Travel Agent) will have a Chatbot interface as a customer service channel to automatically answer some of the customer queries. Since chatbot is generally available on the consumer websites, users are not unfamiliar with such interfaces.

To access Agoda Chatbot, one need to follow below steps as shown in *Figure 1*:

1. Go to Agoda website (<https://www.agoda.com>), or open the Agoda mobile app.
2. Login with your Agoda account.
3. Go to "More" tab.
4. Click "Help Center".
5. Click "Assist me" button.
6. You will be in the Agoda Chatbot interface.

Once the users are on the Chatbot interface, the current servicing flows are pre-defined. There are 8 categories users can select: "Cancellation", "Name Change", "Add guest", "Booking confirmation", "E-receipt", "Date change", "Hotel contact",

"Hotel policies". One can play with the existing customer servicing flows in this Chatbot interface to experience. The Chatbot interface currently has limited functionalities.

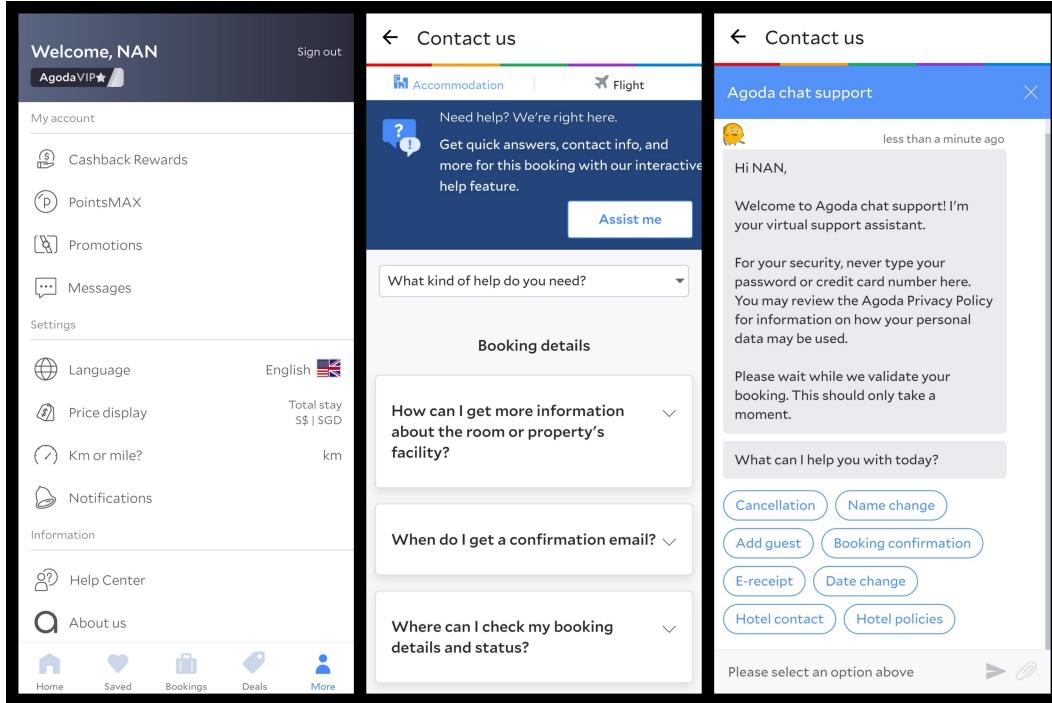


Figure 1—Existing Chatbot Interface of Agoda

2 INITIAL NEEDFINDING

2.1 Needfinding Approaches and Plans

To understand the current weakness of Agoda chatbot interface, **Interview** and **Product Review** are chosen to be methods for needfinding. The reason to use Interview together with Product Review is to reduce the bias. *Table 1* summarizes the potential biases and solutions from both methods.

Also, through product reviews we can gather the general issues of Agoda app, but might not be directly related to the Chatbot interface. We can get more targeted feedback of Chatbot interface from Interviews.

2.1.1 Interview plan

To make surveys more effective, we need to make sure below 5 points (Joyner, 2016c):

Table 1—Needfinding Methods Biases and Solutions

Method	Bias	Reason	Solution
Interview	Social Desirability Bias	People tend to agree with interviewer if the questions have certain tendency	Phrase the question neutrally
Product Review	Voluntary Response Bias	People writing product review usually have strong opinions	We need to use Product Review together with Interview

1. Less is more. Ask minimum questions necessary.
2. Be aware of bias, phrase questions in a neutral way.
3. Tie them to the inventory.
4. Test it out. Have co-workers test survey questions before sending them out.
5. Iterate and revise surveys according to the tests.

Below interview questions will be asked for 10 people to gather the feedback of Chatbot interface of Agoda app.

1. Which app do you use to book hotels when you are travelling?
2. What do you like about Agoda?
3. What do you dislike about Agoda?
4. (After trying the Chatbot interface) What do you like about the interface?
5. (After trying the Chatbot interface) What do you dislike about the interface?

2.1.2 Product Review plan

We will collect and review the top 50 reviews from both Google Play Store and Apple App Store. Because these two stores are the most popular sources for the users to access Agoda app. There are 480K downloads from Play Store and 34.1K ratings from App store. These are the great places to find out what people dislike about the app, especially on the customer service sector. We will use the redesigned Chatbot interface to address those issues.

2.2 Needfinding Conclusion - Results and Summary

2.2.1 Summary and Analysis - Interviews & Product Reviews

For the interview result, all people know the Agoda app, since it is one of the most popular travel apps in the region. The major benefit of using Agoda app is that people can easily book a hotel through the platform. And the major

concern of the app is that the app may not offering the best price. In terms of the Chatbot interface, interviewees generally like the feature that it shows the capabilities upfront. People thus understand what can be done through Chatbot. The major concern of the Chatbot interface is that, most of the scenarios that users need help are not addressed by the existing Chatbot interface. For example, if they want to know if the hotel has a 24hrs reception, or whether it supports quarantine, the current Chatbot interface cannot answer those questions.

For the product reviews, the common things people like are "Website is very easy to use", "User friendly", and "Great value for money". People like the ease of use of existing interface. But seldom do people mention chatbot interface. It may indicate that the entry of chatbot is hard to find. When we look at the common thing people dislike about Agoda, they are "Poor customer service and site issues", "There is no way to get help from Agoda", or "Tried to contact Agoda for help but the BOT directing us was useless". People use the app reviews as a channel for complaints. The major reason of the complaints are due to poor customer service. And people do feel the Chatbot is useless if it cannot address the users' issues and not able to connect the users to real agents in the interface.

2.2.2 Data Inventory - Interviews & Product Reviews

These are the questions we want to answer though our needfinding exercises: (Joyner, 2016d)

- Who are the users?
- Where are the users?
- What is the context of the task?
- What are their goals?
- What do they need?
- What are their tasks?
- What are their subtasks?

Table 2 summarizes data inventories in the Interviews & Product Reviews method.

Table 2—Data Inventory - Interviews & Product Reviews

Who	Where	Context	Goal	Need	Tasks
Users pre-booking	Home	Plan and book hotels	Successfully book hotels on Agoda app	Easily find relevant hotel information through Chatbot interface	Find Chatbot interface, ask questions, follow suggested flow, get answers
Users post-booking	Destination	Need to check-in hotels	Get a receipt for hotel booking	Chatbot to provide steps to get e-receipt	Find Chatbot interface, ask questions, follow suggested flow, get e-receipt
Users post-booking	Destination	Hotel has no room to offer	Cancel booking and find a nearest hotel	Cancel existing booking and make a new booking	Find Chatbot interface, ask questions, follow suggested flow, get real agent chat

3 HEURISTIC EVALUATION

In this section, we will perform the heuristic evaluation to evaluate the existing chatbot interface. We have 15 design principles to evaluate as summarized in *Figure 2*. (Joyner, 2016e) We will discuss areas the current interface meet the design principles and areas it does not.

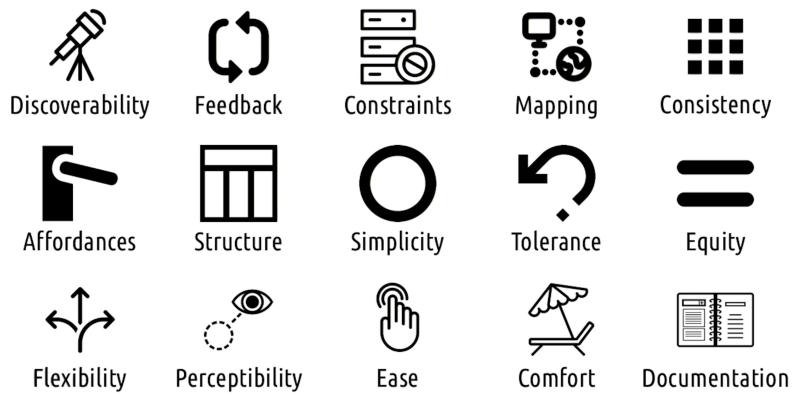


Figure 2—15 Design principles for Heuristic Evaluation

3.1 What works well and What makes it work well

We will summarize what works well and why in the *Table 3*

3.2 What does not work well and Why does not it work well

We will summarize what does not work well and why in the *Table 4*

Table 3—What works well and why for existing chatbot interface

What works well	Design Principle	What makes it work well
Users are well informed about what's going on	Perceptibility	The conversation flow keeps user understand the context, and the user gets fast response
Users will not able to generate wrong input	Constraints	Users can only press predefined intentions buttons
The interface is simple and easy for the user to understand	Simplicity	Interface is clean and with only necessary information
The interface is simple and easy for the user to understand	Ease	Users can click buttons provided and get into the serving flow
Users can use the chatbot interface on mobile or desktop	Comfort	The chatbot interface uses different designs to fit for both mobile and desktop screens
Users speaking different language can all use the chatbot interface	Equity	The chatbot interface support multiple languages

Table 4—What does not work well and why for existing chatbot interface

What does not work well	Design Principle	Why does not it work well
Users are hard to find chatbot interface	Discoverability	To access chatbot interface, one need to go through multiple pages
The normal user input is greyed out and users will be confused by the interface	Affordance	The interface looks like the normal chatting interface, but users are not allowed to key in free text
Users are not able to make mistakes or ask wrong questions	Tolerance	The interface fixes the possible buttons and leave no room for users to make free questions
Users are not able to ask the real issues they are facing other than the predefined 8 intentions	Flexibility	Free text input is disabled by the interface, there is no place to key in
Users will be confusing about what the chatbot interface can do	Consistency	Many other major industry leaders allow user to key in free text and transfer to real agents
The interface looks like a chat interface but it only allows follow the pre-defined flow	Mapping	Users would expect to chat in a chatbot interface, but the interface does not provide that functionality
Users have no instruction or documentation to follow if chatbot fails	Documentation	There is no clear instruction to guide users to real agents if chatbot fails to serve the users
The feedback is not helpful is the chatbot does not understand the input query	Feedback	For the scenarios like chatbot cannot understand users queries for many times, it should transfer to real agent
No structure is used in the current design	Structure	There is no structure information to display in the frontend
Novice users may not need to access chatbot to solve their issues	Expert Blindspot	There should be instructions to guide novice users to read FAQ first before trying the chatbot interface
Users can be frustrated if they are trying to get in touch with real agent	Learned Helplessness	The users will not reach to the real agent no matter what he try, the interface does not provide an "exit" option for the user using chatbot

4 INTERFACE REDESIGN

In this section, we will redesign the chatbot interface using **high-fidelity wireframe prototype**. The interface will be designed to address the data inventories in section 2. Comparing to the existing interface, the redesign will allow users to key in free text to get the answers. Also, the new interface will auto-complete the user inputs, like when you are typing questions in Google. If the Chatbot knows the user intents, it will guide the user to get the answers. If not, the Chatbot will be transferred to real agent after failing to serve the customer for 3 times. The overall user journey in the redesigned chatbot interface is shown in the *Figure 3*

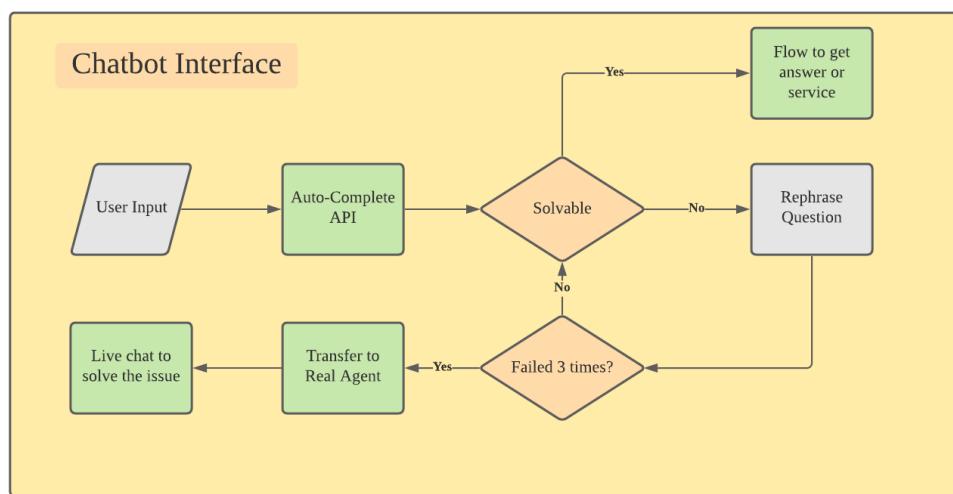


Figure 3—Overall Flowchart of the redesigned chatbot interface

4.1 User Journey of Solvable Issue on Chatbot Interface

As shown in *Figure 3*, users would be able to get the desired answers through the redesigned chatbot interface. The interface will have an auto-complete function like Google to help the user finish the question. Then the interface will suggest users to follow steps to have the information they need.

4.2 User Journey of Unsolvable Issue on Chatbot Interface

As shown in *Figure 4*, users will be asked to rephrase the question if the Chatbot cannot understand the intent. After 3 times failure, the Chatbot may not have the capability to understand the complex user query, then the user would be able to connect to the real agent to solve the issue.

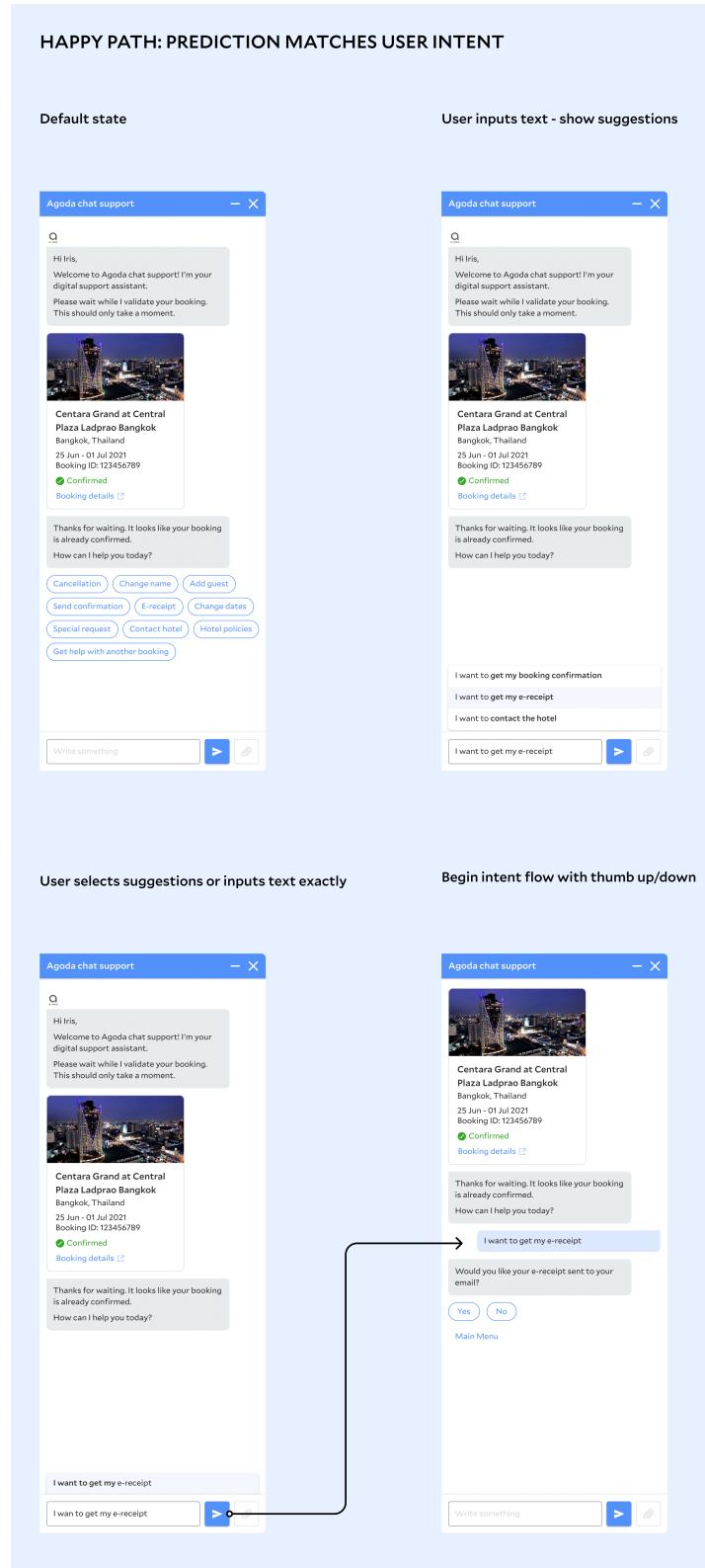


Figure 4—User journey to get e-receipt for hotel booking

UNHAPPY PATH: PREDICTION DOES NOT MATCH USER INTENT

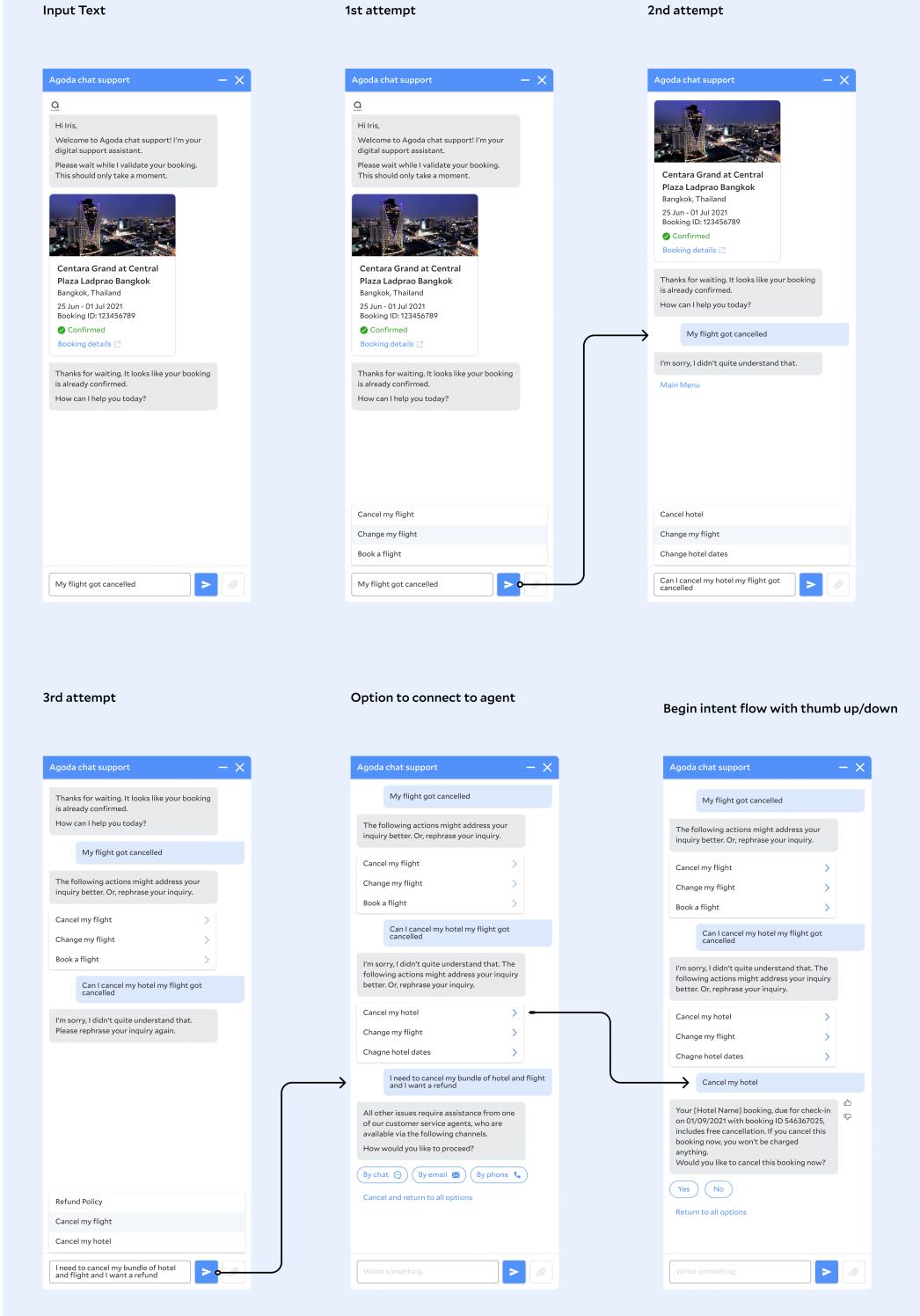


Figure 5—User journey to get e-receipt for hotel booking

5 INTERFACE JUSTIFICATION

From the redesign of the chatbot interface (at the end of this report), we can see that the criticisms from first section is addressed while preserving the positive elements from the original interface. We have enable the user to key in free text, and help users to auto-complete sentence if they are not entirely clear about what questions to ask. Also, if the users fail many times by chatbot, the interface will show an option to connect to the real agent.

5.1 How redesigned interface addresses the requirements from Needfinding

Through the redesign, the interface now shorten the gulfs of execution and gulfs of evaluation as shown in *Table 5*. The original interface is using the processor view, it assumes 8 pre-defined scenarios for the users (as shown in the *Figure 6*). But the redesigned interface is using the participant view, it takes the user context into consideration that it will allow users to connect to real agent if the user query is unique to the context and not able to be addressed by the chatbot itself.

Table 5—Gulfs of execution and evaluation

Name	Stage 1	Stage 2	Stage 3
Gulfs of execution	Identify Intentions	Identify Actions	Execute in Interface
Gulfs of evaluation	Interface Output	Interpretation	Evaluation

	Processor	Predictor	Participant
Requirement	Fit within human limits	Fit within user's knowledge	Fit within usage context
Evaluation	Quantitative experiments	Qualitative studies	<i>In situ</i> (real-world) studies
Pros	Objective comparisons May use existing data	Fuller picture May target novices	Authentic context Faithful user attention
Cons	Expert-focused What, not why Optimize, not redesign	Expensive to analyze Prone to biases Ignores context	Expensive to perform Needs real interfaces Subject to interference

Figure 6—Processor View, Predictor View, and Participant View

5.2 Summary Table - Issues, Solutions and Design Principles

Table 6 summarize all the improvements to address the design principle issues in current interface.

Table 6—Issues, Solutions in Redesign and Design Principles

Original Issues	Design Principles	Redesign Solutions
Users are hard to find chatbot interface	Discoverability	We will move the access of chatbot to the home page, users can do one-click access
The normal user input is greyed out and users will be confused by the interface	Affordance	The interface allows free text, users can use the text input box as it is
Users are not able to make mistakes or ask wrong questions	Tolerance	The interface will allow users to ask any questions, and transfer to real agent if not able to handle the query
Users are not able to ask the real issues they are facing other than the predefined 8 intentions	Flexibility	Free text input is enabled by the interface, users are allowed to ask all kinds of questions
Users will be confusing about what the chatbot interface can do	Consistency	The chatbot interface will have the consistency as other chatbot interfaces to reduce the cognitive effort of users
The interface looks like a chat interface but it only allows follow the pre-defined flow	Mapping	Chatbot interface will be similar to day-to-day chat interface like WhatsApp or Telegram, it will bridge the gulfs of execution
The feedback is not helpful is the chatbot does not understand the input query	Feedback	The chatbot will give feedback such as "I don't understand", and transfer to real agent after 2 times failure
Novice users may not need to access chatbot to solve their issues	Expert Blindspot	A first-timer walk-through will pop up if the user is logging in for the first time.
Users can be frustrated if they are trying to get in touch with real agent	Learned Helplessness	The interface will provide an "exit" option for the user using chatbot and cannot find the answer after many times trying

6 EVALUATION PLAN

According to Joyner, *Table 7* are the advantages for different methods: (Joyner, 2016a)

Table 7—Advantages for different evaluation methods (Joyner, 2016a)

Advantages	Qualitative	Empirical	Predictive
Informs ongoing design decisions	Yes	-	Yes
Investigates the participant's thought process	Yes	-	Yes
Draws conclusions from actual participants	Yes	Yes	-
Identifies provable advantages	-	Yes	-
Provides generalizable conclusions	-	Yes	-
Does not require any actual users	-	-	Yes

Since the chatbot interface we are studying is an existing interface and we would like to evaluate the performance of the redesigned interface, we will use empirical evaluation to evaluate.

6.1 Empirical Evaluation

Control Group: Randomly assigned people, 50% of the total test population.

Experimental Group: Randomly assigned people, the rest 50% of the total test population.

Null Hypotheses H₀: The redesigned chatbot interface is not significantly better than the original interface.

Alternative Hypotheses H₁: The redesigned chatbot interface is significantly better than the original interface.

Independent Variables: The user demographics, the device types, the Geo-locations, dates and all other Agoda app interfaces shall be independent variables that both control group and experimental group have same distribution.

Dependent Variables: The only difference between control group and experimental group is the Chatbot interface of Agoda app. So the only dependent variable shall be the chatbot interface.

We shall conduct a **between-subjects** study, daily visitors of Agoda app are

randomly divided into control group (50%) and experimental group (50%) to perform the **A/B testing**. Since Agoda currently has 75 million active users, a few days of experimentation will give us enough confidence about the outcome. The outcome signal is based on the user interaction by the end of the chatbot flow, whether the user select "Yes" or "No" to the final question "Is the issue solved?". We will exam the statistical significance (p-value) after the experiment to decide whether to reject null hypothesis.

6.2 Planned Statistical Analysis

We can refer to *Table 8* to select the proper empirical test for the redesigned chatbot interface. Because in this experiment, we have categorical independent variables and binomial dependent variable, we shall use **Binomial test** to evaluate. If the final result is statistically significant (**p-value <= 0.05**), we can reject the null hypothesis and conclude that the redesigned chatbot interface is significantly better than the original interface.

Table 8—Empirical Tests Selection (Joyner, 2016b)

IV	DV	Treatments	Recommended
Categorical	Nominal	2 or more	Chi-Squared test
Categorical	Ordinal	2	KS test
Categorical	Interval or Ratio	2	Student's t-test
Categorical	Ordinal	3 or more	Chi-Squared test
Categorical	Interval or Ratio	3 or more	ANOVA
Categorical	Binomial	1 or 2	Binomial test
Interval or Ratio	Interval or Ratio	2	Linear Regression

7 REFERENCES

- [1] Joyner, David (2016c). *5 Tips: Interviews*. URL: <https://classroom.udacity.com/courses/ud400/lessons/9662642568/concepts/97510951660923>. (accessed: 30.07.2021).
- [2] Joyner, David (2016d). *Data Inventory*. URL: <https://classroom.udacity.com/courses/ud400/lessons/9662642568/concepts/96550826820923>. (accessed: 30.07.2021).
- [3] Joyner, David (2016a). *Exercise: Evaluation Pros and Cons*. URL: <https://classroom.udacity.com/courses/ud400/lessons/9753906972/concepts/97594129880923>. (accessed: 30.07.2021).
- [4] Joyner, David (2016e). *Other Sets of Principles*. URL: <https://classroom.udacity.com/courses/ud400/lessons/9296683706/concepts/92967534060923>. (accessed: 30.07.2021).
- [5] Joyner, David (2016b). *Quantitative Data and Empirical Tests*. URL: <https://classroom.udacity.com/courses/ud400/lessons/9753906972/concepts/5b6619b7-b1e3-47ac-beac-f5d5377d3151>. (accessed: 30.07.2021).