Evaluation Protocol & Next Steps

Telempathy (Usability Testing)

Group 1 — Xin An, Joshua Shum, & Jinghui Yao
INF1602H — Fundamentals of UX

1. Goals

1. To determine if *Telempathy* is easy to use and intuitive to learn.
2. To determine whether or not *Telempathy* can facilitate the access of mental health resources and services within five minutes.
3. To determine whether or not *Telempathy* is a viable or preferred medium of communication for users.

2. Questions

1. To determine if *Telempathy* is easy to use and intuitive to learn.
    a. Is the app easy and intuitive to navigate?
    b. Are the intended functions of the main features (e.g. Navigator, Emergency Services, Peer Stories, & Resource Library) easy to understand?
    c. Do users experience any confusion about using any of the features?
    d. Currently, we have a help manual accessible in the hamburger menu of the app. Is it preferable to add an information/help button for specific features?
    e. Do users hesitate or make errors when navigating *Telempathy*?
2. To determine whether or not *Telempathy* can facilitate the access of mental health resources and services within five minutes.
    a. Does it take users longer than five minutes to access any mental health resource or service (e.g. any of the main features)?
    b. Does it take users longer than five minutes before *Telempathy* offers personalized suggestions?
    c. Is the “first available appointment” prediction and booking option useful?
3. To determine whether or not *Telempathy* is a viable or preferred medium of communication for users.
    a. Are there pain points or design functionalities in the interface that may deter users from using the app?
        i. As this is a medium-fidelity Balsamiq wireframe prototype, it does not actually have access to all of the University of Toronto’s resources and services. As such, this question is based on the notion of whether or not any UI/UX design problems can be identified through usability testing.
    b. Is a mobile app a more preferred method of finding and accessing these mental health services (as opposed to existing forms, e.g. email, Student Life app, Student Life website)?
    c. Does *Telempathy* allow users to back out whenever they want?
    d. Does *Telempathy* satisfy the privacy and safety concerns of users?

3. Evaluation Methods

3.1. Overview

We chose *usability testing* and post-observation *semi-structured interviews* as our evaluation methods for *Telempathy*'s clickable medium-fidelity prototype. The data collected from these methods will likely complement each other and elicit a broader understanding of whether the user experience and usability of *Telempathy* meets our aforementioned goals.

3.2. Usability Testing (User Observations)

This was selected so we could record metrics such as:

- Task success (Pass/Fail, or 0-100%)
- Time elapsed
- Task efficiency
- Observed instances of user errors or hesitation

3.3. Semi-Structured Interviews

This was selected to elicit:

- Positive feedback (What worked? What did users like?)
- Negative feedback (What didn't work? What did users dislike?)
- Suggestions and ideas for change or implementation
- How likely would users be willing to use this app in the future?
    - Compared to other methods?
    - Overall?

4. Identifying Practical Issues

4.1. Overview

Three tasks were designed to ensure that many of our questions and concerns within our goals could be addressed. However, as the two main features of the apps were a *personalized mental health care navigator* and an *immediate help/emergency button*, the tasks we designed for participants to carry out are typical in what we would expect of Lee (our persona). These tasks address all of our hills, in that:

1. Lee/UofT students will be able to find and access available resources through **one centralized platform** (Tasks 2 & 3)
2. Lee/UofT students will be able to speak with someone or something they can trust, with campus-specific expertise, **at any time** throughout the day or week (Tasks 1, 2, & 3)
3. Lee/UofT students will be able to get personalized suggestions for resources and services in less than **5 minutes** of initiating their search (Tasks 2 & 3)

A mock run of the test was conducted by all members of the design team to identify any problems that may come up during usability testing. Through this, we identified several issues that would need to be addressed as extra instructions to participants prior to testing, in order to prevent introducing bias or reliability concerns to the study:

1. Due to the limited functions of Balsamiq, not all interfaces are fully clickable. Certain test elements are thus predetermined and non-interactive, such as:
    a. Options 1, Options 2, and Options 3: Option selection (opening and closing drop-down list for more details will not work)
    b. Booking Calendar: Date selection (Balsamiq dynamic object that detects the date and changes the object accordingly)
    c. Booking Calendar (Pop-Up): Time slot selection
2. As the resource directory and navigator does not actually have live access to all of the University of Toronto's resources, options provided on screens are predetermined and are not reflective of the intended functionality of the app.

4.2. Tasks

TASKS	CORRECT STEPS
Task 1. Call an emergency service (with or without logging in)	**Without Login:** **1.** From Login screen, select EMERGENCY **2.** Tap to call any emergency service **3.** Arrive at dialing screen **Login:** **1.** From Login screen, select LOG IN **2.** Select EMERGENCY SERVICES **3.** Tap to call any emergency service **4.** Arrive at dialing screen
Task 2. Using the Navigator, find a resource/service called "My SSP for International Students", then book an appointment with the time recommended by *Telempathy*.	**1.** From Login screen, select LOG IN **2.** Select YOUR NAVIGATOR **3.** Tap "send" icon or "mic" icon) **4.** Tap YES **5.** Select NEXT **6.** Select NEXT **7.** Select My SSP for International Students and tap MORE DETAILS **8.** Select BOOK **9.** Visually confirm details and select CONFIRM **10.** Arrive at Booking Confirmation (Success)
Task 3. Using the Navigator, book an appointment with any option for [today], from 11:00 AM — 12:00 PM	**1.** From Login screen, select LOG IN[1] **2.** Select YOUR NAVIGATOR **3.** Tap "send" icon or "mic" icon) **4.** Tap YES **5.** Select any option and tap MORE DETAILS **6.** Select CALENDAR **7.** Tap on today's date and press SELECT[2] **8.** Tap on the "11:00 AM — 12:00 PM" time slot and press BOOK[3] **9.** Visually confirm details and select CONFIRM **10.** Arrive at Booking Confirmation (Success) screen

[1] Alternate Step 1: From previous success screen, select RETURN TO HOME PAGE
[2] In the test prototype, the calendar's date is dynamic and pre-selected. If the user does not tap specifically tap on today's date, this will not be counted as an error.
[3] In the test prototype, the time slot "11:00 AM — 12:00 PM" is pre-selected. If the user selects another time and presses CONFIRM, this will not be counted as an error.

4.3. Post-Observation Interview Questions

1. What are your initial thoughts of *Telempathy*?
2. Was it straightforward and easy to use? Why or why not?
3. What did you like about it?
4. What did you not like about it?
5. What would you change?
6. Do you have any ideas or suggestions to improve *Telempathy*?
7. If this app was developed, would you use this app? Why or why not?
8. Do you have any other questions or comments?

4.3. User Selection

Three University of Toronto students were recruited to participate in the usability testing and interviews based on affirmative answers to the following screener questions:

1. Are you currently 18 years old or more?
2. Are you a current University of Toronto student?
3. Do you own a mobile device with internet access?

4.4. Testing Conditions

To best emulate the conditions in which a typical representative user is most likely to use the app, we allowed participants to select a location of their choice based on secludedness, personal level of comfort, and a low level of noise.[4]

To emulate a real device, we used Balsamiq's User Testing presentation mode by opening the URL to the prototype on a mobile browser in full-screen. This allowed participants to use device interfaces to navigate *Telempathy* while simultaneously preventing them from accidentally navigating to other screens to prevent confusion.

4.5. Evaluation Process

To ensure that our data is accurately recorded during testing, Ann will act as the moderator who gives an explanation of *Telempathy,* any additional instructions, and the tasks as the tests proceeds, whereas Vicky will be the observer and taking notes of what the participant says, any questions that come up, as well as user hesitations or errors.

Participants will then be asked the post-observation interview questions by Ann, while Vicky records their answers by typing notes using Google Suite (Docs or Sheets).

[4] The assumption is that *Telempathy* users are most likely to seek mental health resources through our app at a quiet and comfortable location.

5. Ethical Concerns

5.1. Overview

Throughout this evaluation, participant confidentiality will be maintained, and personally identifiable information will be discarded. Participants will be anonymized through assigned participant identification numbers (P1, P2, P3, etc.). No demographic information will be collected or used other than for screening purposes.

Participants were informed about the purpose of the study beforehand, how the collected data would be used, and were able to withdraw at any time without penalty. No compensation was offered for their participation or time.

5.2. Pre-Interview Script

Hi, my name is **<name>**, I am a graduate student under the direction of Professor Olivier St-Cyr in the Faculty of Information, University of Toronto. For a course titled Fundamentals of User-Experience (INF1602), I am conducting a usability test study for the purpose of designing an app that helps students find and access mental health resources and services at the University of Toronto.

Your participation in this interview is voluntary and will not be compensated. If you choose not to participate or to withdraw from the interview at any time, there will be no penalty. During the research process, another individual from our design team, **<name>**, will be observing and taking notes. Your name will not be recorded and all identifying personal information will be removed during our data analysis. Instead, you will be assigned a participant identification number (such as P1). If you have any questions concerning this, you may email me at **<email here>**.

There are no right or wrong answers. The goal of our usability test and interview is to gather information about your experience with our app. Please verbalize any thoughts and questions that may come up throughout the usability test, and take your time thinking and responding throughout the interview.

Do you have any questions? If not, let’s begin!

6. Data Evaluation, Analysis, & Interpretation

6.1. Usability Testing Data

6.1.1. Task #1 (Success/Fail)

- P1: Complete success, without assistance
  - ↳ User used login path
- P2: Complete success, without assistance
  - ↳ User used no-login path
- P3: Complete success, without assistance
  - ↳ User used no-login path

6.1.2. Task #2 (Success/Fail)

- P1: Complete success, without assistance
- P2: Partial success, without assistance
  - ↳ User error occurred due to confusion regarding button functions
- P3: Partial failure, without assistance
  - ↳ User gave up; failed to see third page of options
  - ↳ User expressed confusion regarding some button functions (same as P2)

6.1.3. Task #3 (Success/Fail)

- P1: Partial success, without assistance
  - ↳ User navigated elsewhere before arriving at the correct screen
- P2: Complete success, without assistance
- P3: Complete success, without assistance

6.1.4. Time Elapsed (MM:SS)

- **By user (Average → 00:36)**
  - ↳ P1 (Average → 00:38)
    - Task 1: 00:18
    - Task 2: 00:35
    - Task 3: 01:00
  - ↳ P2 (Average → 00:30)
    - Task 1: 00:15
    - Task 2: 00:53
    - Task 3: 00:22
  - ↳ P3 (Average → 00:41)
    - Task 1: 00:13
    - Task 2: 01:11
    - Task 3: 00:39
- **By task (Average → 00:36)**
  - ↳ Task 1: 00:15
  - ↳ Task 2: 00:53
  - ↳ Task 3: 00:40

6.1.5. Efficiency (Steps/Clicks Required)

- Task #1
  - ↳ All users completed the task within the correct and exact number of steps required (2-3 steps).
- Task #2
  - ↳ Users took between 9 and 12 steps. Correct number of steps was 9.
  - ↳ P1 → 9 steps. P2 → 12 steps. P3 → 10 steps.
- Task #3
  - ↳ Users took between 9 and 19 steps. Correct number of steps was 9.
  - ↳ P1 → 19 steps. P2 → 9 steps. P3 → 10 steps.

6.1.6. Errors & Issues

- Task #1
  - ↳ All users completed this task successfully with no errors or issues.
- Task #2
  - ↳ P1 completed this task with no errors or issues.
  - ↳ P2 and P3 partially succeeded. Users expressed that they found the intended function of Navigator unclear from the descriptive text on the home page.
  - ↳ Users expressed confusion over the “Next” button on the options pages.
- Task #3
  - ↳ All users completed this task successfully.
  - ↳ Users expressed confusion over the “Next” button on the options pages.

6.2. Interview Data

6.2.1. Positive Feedback

- “I like the navigator. It is cool to talk with an AI. Also the log in page looks beautiful. What's more, I think it is good to have 'remember me'. You know, there's no such a thing in the Acorn and Quercus that I need to type my account every time. So annoying.” (P2)
- “The emergency [button] is very helpful” (P3)
- “I like the function page. It only includes four features, simple and clear. Won't be overwhelmed by information. Also, the emergency button is very convenient.” (P3)

6.2.2. Negative Feedback

- “I thought after I clicked the 'Yes', I will see a page with all the options available. I don't know what 'Next' can do. The 'page 1 of 3' doesn't make sense to me and it is very small.” (P1)
- “How can I know that I can make an appointment in the navigator. The short description under the 'Navigator' doesn't imply that I can make an appointment here.” (P2)
- “For the peer stories part, I'm not sure if students are willing to share their stories since everything related to mental health could be sensitive.” (P3)

6.2.3. Suggestions & Ideas

- “I think it would be better to list all options in one page and then when clicking each option, you can see the introduction of the option.” (P1)
- “Maybe there should be a search function so I can search for a keyword like 'reservation'?” (P2)
- “It would be useful and time-saving if there are some options/keywords I can choose when talking to the navigator.” (P3)
- "I think it would be better if all the options are in one page and we could scroll down to see all options, then I won't miss anything.” (P3)

6.2.4. Additional Questions & Comments

- “Usually I would go to the wellness center or call the doctor. I am willing to call the doctor so I will search for the phone number or the address of the service. Talking to a real person makes me feel it is reliable.” (P1)
- “So can I see details of options in the resource library?” (P3)
- “What is the “Options” page about? (P3)
- “Why can't I click the services through the 'Resource Library'? So I must use the 'Navigator' to access the services?” (P3)

6.2.5. “Would you use this app?”

- “I don't think the app is useful. Since I think students needed for mental health services would think an AI is not reliable. There is no real person to answer your question in this app.” (P1)
- “I think students will need this app since there is no similar product on the market currently.” (P2)
- “Maybe the App is good but because of my bad experience with AI, I would like to use Google.” (P3)

6.3. Methodological Reflections

6.3.1. Reliability

Our usability tests involved a set number of three tasks to be conducted with users, with a specific order of correct steps that users would have to take in order to be considered successful in completing the task. Alternate “success” paths and criteria were also identified in the design of the tasks, as some of our tasks are based on the fulfillment of the users’ information needs or personal level of satisfaction.

Our post-observation, semi-structured interview consisted of eight open-ended, predefined questions that allowed us to elicit constructive criticism, feedback, and suggestions to base our next steps on.

Given that we used the same set of mediators and observers for all participants, we were able to ensure that the methodological process and data collection was controlled and consistent, thus avoiding any unexpected or missing data points which may have affected the results. As such, we feel that our results are reliable, consistent, and could be replicated with other participants.

However, it is necessary to address the one questionable metric — the “Time Elapsed” measurement. As aforementioned, Balsamiq has limitations that did not allow

us to create a fully clickable prototype. As such, users were not offered the opportunity to read more in-depth details (e.g. costs, specific information about services, hours) about chosen options that would likely influence their decision-making. In real life, users would likely need more time to read about options before making a decision. In such a way, the "Time Elapsed" section does not necessarily reflect the time it takes for a user to access information that he or she considers satisfactory for their information-seeking needs, but only the time it takes for the user to open the app and access information

However, as the goal of *Telempathy* was to "facilitate the access of resources and services within five minutes" (as referenced in Goal #2) rather than the *fulfillment* of the user's information needs, we still consider this metric a reliable indicator that we have been successful in this goal.

6.3.2. Avoiding Bias

The same two members participate in testing allowed us to deliver consistent methodological procedures, and ensure that all data points were comparable.

Our involvement was limited to one mediator and one observer. The mediator's role was to introduce ourselves and *Telempathy*, explain the reasons for our testing, read the privacy/ethical concerns script, deliver the extra instructions as identified in Section 4.1., prior to the start of the testing, and give users tasks to complete. The observer's task was to gather and document data, such that the mediator would be able to address potential user questions without having to worry about missing observations and affecting the overall time elapsed.

Furthermore, our usability testing team only answered points of clarification regarding the tasks' objectives during the usability tests. No assistance or personal comments were rendered unless directly asked by participants in extraordinary or unexpected circumstances (e.g. for language translation or clarification).

6.3.3. Testing Environment

In order to create a testing environment that emulated real life situations, we designed our usability tests to be conducted on a mobile device at a location selected by the participant. As our research suggested that users more frequently seek mental health help when comfortable, it was necessary to select a location considered comfortable to participants. Furthermore, using a mobile device was the most appropriate method to assess *Telempathy*'s usability: as our project is a mobile app, using a mobile device as part of the testing environment was more preferable than with a laptop or personal computer.

However, we experienced some difficulty with opening the Balsamiq prototype due to browser incompatibilities or issues, so for some participants, usability testing was conducted on the mediator's personal device (laptop) rather than a mobile device.

6.3.4. Scope

A total of three tasks were each conducted with three participants in order to get a sense of the current functionality of *Telempathy* from users outside of our design team and INF1602. This allowed us to get a fresh perspective of *Telempathy* from our intended users. Limiting our participant involvement to three also allowed us more time/freedom to have in-depth discussions to inform our next design objectives.

Given that only a minimal viable product (MVP) has been developed at this point, a scope of three participants at this stage is appropriate as to get a clear sense as to the elements that should be immediately focused on for development. In future iterations, additional tasks would be designed to test other features of *Telempathy*.

6.4. Summary (Revisiting Goals & Questions)

6.4.1. Revisiting Goal #1

Goal #1: To determine if *Telempathy* is easy to use and intuitive to learn.

Based on our data, we can confidently say that our interface is easy to use and intuitive to learn for Task 1. However, for Task 2 and 3, users expressed confusion in regards to several of the button labels and functions. As such, despite the relatively high success rate (8 of 9 tasks), we think there is room for improvement in regards to this goal.

6.4.2. Revisiting Goal #2

Goal #2: To determine whether or not *Telempathy* can facilitate the access of mental health resources and services within five minutes.

Our data suggests that despite confusion, users can easily access mental health resources and services using *Telempathy*, well within five minutes. The average task completion time was 36 seconds.

However, as aforementioned, in real life, users would likely need more time to read about options before making a decision. In such a way, this metric only measures the time it takes for the user to open the app and access information — and does not account for the time spent by the user reading through information and deciding. Nonetheless, this is a huge improvement compared to existing solutions (as tested through usability testing in the user research phase)

6.4.2. Revisiting Goal #3

Goal #3: To determine whether or not *Telempathy* is a viable or preferred medium of communication for users.

Data in this area was inconclusive. One user found the concept of a mental health AI “cool” and that *Telempathy* would be useful as there are no similar apps available in the market. However, others expressed hesitation due to previous bad experiences with AIs — suggesting that they would rather call or go in-person. More user research with a higher number of participants would be required to confidently address this goal.

7. Next Steps

1. Adjust descriptions on “Home Page” to clarify purpose of features
    a. Usability testing suggested that our descriptions on the Home Page did not explicitly state that users could book an appointment through the Navigator or Resource Library features.
    b. To address this, we will adjust our descriptions with clearer wording
2. Continue development of “Peer Stories” and “Resource Library” features to provide more methods of accessing resources
3. Completely re-examine our approach for displaying information on the Navigator Option pages
    a. Change label on Navigator Option pages from “Next” to “More” to clarify the button’s function
    b. Implement search/filtering function within Navigator Options to allow users to narrow/refine list of available options
    c. Two users suggested one page of options with a scrollable list, but this may not be user-friendly as the list would not be ranked/sorted in any particular way, and has no way of refining this list
        i. This might induce information overload as there might be so many options that it can overwhelm the user
4. Conduct further research with more participants to determine if there is a viable market for this sort of app