Empathetic Conversational Agent Assisting with Title IX Cases and Digital Counseling

Brandon Persons
Advisor: Professor Soussan Djamasi
Worcester Polytechnic Institute, 2020
# Table of Contents

Abstract ........................................................................................................................................ 4
Introduction ................................................................................................................................... 5
Background ................................................................................................................................... 6

- **ERIN: A Friendly Virtual Assistant** ................................................................................... 6
- **Human-Chatbot Interaction** ................................................................................................. 7
- **Entities and Intents** ............................................................................................................. 7
- **Personas** .............................................................................................................................. 8
- **Current Solution** .................................................................................................................. 8
- **Opportunities for Improving the Current Solution** ............................................................... 9

Developing the Chatbot ............................................................................................................... 9

- **Target Users** ....................................................................................................................... 9
- **Infrastructure** ..................................................................................................................... 13
- **Interface** ............................................................................................................................. 15
- **Conversational Scripts** ....................................................................................................... 16

User Studies ................................................................................................................................ 17

Study 1 – User Interviews............................................................................................................ 17

- **Overview and Design** ......................................................................................................... 17
- **Participants** ......................................................................................................................... 18
- **Measurement** ..................................................................................................................... 18
- **Results** ............................................................................................................................... 19

Study 2 – User Experience and Usability Testing .................................................................... 23

- **Overview** ........................................................................................................................... 23
- **Participants and Design** ...................................................................................................... 24
- **Measurement** ..................................................................................................................... 25
- **Results** ............................................................................................................................... 26

Discussion ................................................................................................................................... 29

Limitations ................................................................................................................................... 31

Future Work .................................................................................................................................. 31
Contribution .................................................................................................................................32
  Theoretical ..................................................................................................................................32
  Practical .......................................................................................................................................33
Bibliography ....................................................................................................................................35
Appendices .....................................................................................................................................41
  Appendix A – ERIN .......................................................................................................................41
  Appendix B – Design Iterations ....................................................................................................42
  Appendix C – Conversational Scripts ...........................................................................................43
  Appendix D – Personas ................................................................................................................45
Abstract:

Based on a recent report by the Office for Civil Rights in 2020, Title IX and mental health issues have been increasing across all college campuses through the United States. Because of this increasing trend, many Student Development and Counseling Centers are searching for ways to simplify the process of obtaining resources for college students (Morse, 2020). This project discusses a user-centered design process utilized to develop an application for such services. To address this need by a user-friendly product, this IQP sponsored by the User Experience and Decision Making (UXDM) lab at WPI developed the Empathetic Research IoT Network (ERIN) chatbot, which is seen as an empathetic software agent that provides resources based on users’ need. The goal of this project was to develop and test the first prototype for the ERIN chatbot. The project was started with key informant interviews for proto-persona development. These interviews also provided requirements for software development. After developing the first prototype, user tests were conducted to verify the developed personas and gather user reactions to the first prototype. The user tests showed favorable usability scores for the chatbot. In particular, users reported a more favorable usability rating when they interacted with the ERIN chatbot on their mobile phone.

Keywords: User Experience, Usability, Chatbot, Design, Development, Persona, Natural Language Processing, User-Centered, Human-Chatbot Interaction, Utterances, Entities, Intents, Conversational User Interface, Mental Health, Title IX
Introduction:

As of May 29, 2020, there are currently 1,418 pending cases related to all Title IX categories across the United States (Office for Civil Rights, 2020). This doesn’t include the countless number of closed cases throughout the years since Title IX was enacted in 1972. Additionally, 17.7 million people going through depressive episodes and 48 million people coping with anxiety annually reported by the National Alliance on Mental Illness (NAMI) in 2019. Anxiety and depression are the top two mental health illness seen in young adults and eventually can lead to health concerns if not properly handled (NAMI, 2019). These health concerns and other consequences that occur from the two conditions can lead to suicidal thoughts or actions. Suicide is the second leading cause of death among young adults in the U.S. (NAMI, 2019). Title IX cases and mental health related issues are present on all college campuses affecting many students one way or another. Many students suffer from anxiety and depression because of increased workload (e.g. assignments and exams) and intense social interactions (e.g. student clubs and festivities) which are part of the college experiences. It is important to help students to access information that can help them overcome their problems in a simple way (Morse, 2020). To address this need, this IQP creates a centralized hub for obtaining information related to mental health in an efficient, empathetic, and conversational way through a chatbot. Using the user experience driven innovation (UXDI) paradigm (Djamasbi and Strong, 2019), this chatbot was developed an experience-first approach. That is, the chatbot design was informed by gaining a deeper understanding of user needs and utilized many user experience (UX) and user-centered design techniques such as persona development and user research. These techniques allowed for the development of a chatbot that not only meets the requirements of the project but
also can be expanded to include other user populations and/or address needs other than access to mental health resources.

The paper is organized in a chronological order explaining the processes and information used through the duration of the project. The first part of the paper explains the background information related to the project such as the virtual assistant named ERIN, human-chatbot interaction, and how the user’s intent relates to the functionality of the chatbot (e.g. entities). Additionally, the current solution on WPI’s campus is discussed. This discussion includes problems that were determined through key informant interviews. In the next part of the paper, the development process of the chatbot is presented for each aspect from the conversational scripts to the user-centered design process that utilized proto personas. Then, the testing procedures and results of the user tests are reported. Finally, the limitations and the future development of the chatbot are discussed.

**Background:**

**ERIN: A Friendly Virtual Assistant**

ERIN was developed at the UXDM lab at WPI based upon the premise of creating a virtual assistant for college students. Virtual assistants are developed to help users with common problems in a variety of domains (Bodart and Condon, 2005). There are many commercially available virtual assistants to this date such as SIRI, Cortana, and Alexa, so the goal of ERIN was to innovate the current solutions by incorporating useful applications that have not been utilized in virtual assistants before. To achieve this goal, applications such as emotion recognition and don’t touch your face (an application for minimizing the chances of a COVID infection) were implemented into ERIN. The modular design of ERIN was planned to facilitate flexibility and growth; many more applications will be implemented into ERIN’s infrastructure to make it a
more desirable assistant to all college students, as well as other organizations in the future. Appendix A provides a screenshot of ERIN’s login screen and URL.

**Human-Chatbot Interaction**

Interest in chatbots have been increasing throughout the years due to their desirable implementation for completing tasks in a productive, efficient, and helpful manner (Brandtzaeg and Folstad, 2017). Chatbot’s have been utilized by many corporate companies such as Royal Bank of Scotland, Disney, and Domino’s, in order to complete a service rather than having a difficult to navigate website with FAQs. These companies utilize chatbots since they provide a more natural user interface allowing users to talk to a machine rather than just staying on a neutral, expressionless webpage (Park et al., 2018). This is the premise of human-chatbot interaction because it incorporates conversational user interfaces that users can use as an interactive experience while achieving their goal of using the application. Human-chatbot interaction allows users to have a conversation like they are talking to a friend which is a natural conversational situation for many people (Park et al., 2018).

**Entities and Intents**

Natural language processing (NLP) takes an utterance (e.g. a sentence typed or said by the user) and extracts the intention of the user to determine what the chatbot should return (Singh, 2017). NLP can be broken down into intents and entities. Intents are categorized by two separate categories: casual intents and business intents. Casual intents are characterized as an opening or closing statement of a conversation such as hello or goodbye, where as a business intent maps directly to the service of the chatbot is being utilized for (Singh, 2017). The entity is what the user wants to receive from using the chatbot. Entities are attached to intents (specifically business intents) and are used to obtain the desired outcome from the data that the
chatbot has access to (Singh, 2017). For example, if a user would like the year a movie was released and asked the chatbot for it, the entity in this case would be a date the movie was released. Intents and entities are used to make the chatbot dynamic and allow for chatbots to be less defined by the predefined responses that they are coded with (Cameron et al., 2018).

**Personas**

Personas are developed to give a “fictional, specific, concrete representation of the target users” (Miaskiewicz and Kozar, 2011; Jain et al., 2019). When developing a user-centered product, it is important to understand the user and his/her needs. Personas provide the ability to focus in on the target users and give project stakeholders (e.g. developers, designers, managers) a consistent, clear view of who the product is being developed for (Pruitt et al., 2003; Jain et al. 2019). Personas can be grouped into two major categories: proto personas and research personas. Proto personas, which provide a quick view of the target users, are developed without directly interacting with the user, they are typically developed through assumptions or knowledge about the target user. Research personas are developed through direct interaction with the user (Jain et al. 2019). Both proto and research personas are powerful tools for developing new products and/or improving existing solutions (Miaskiewicz and Kozar, 2011; Jain et al., 2019).

**Current Solution**

Currently at WPI, resources are displayed within the main WPI website in subpages (e.g. Student Counseling and Development Center and Title IX). Each subpage has their own pages for resources. These pages are accessed through a search bar or a navigation bar on the side of the screen. The main color scheme of the website is crimson red, black, and white, which is consistent with the color scheme of other WPI webpages.
Opportunities for Improving the Current Solution

The interview with the Director of Counseling at WPI revealed that the complex structure of the website could make it difficult for students to get access to the resources they desired. It is particularly important to minimize cognitive effort for students when they are dealing with a Title IX or mental health issue, due to their heightened emotions. Additionally, the red color of the website could potentially trigger or further intensify heightened emotions (Morse, 2020). The director desired a welcoming, easy to access, centralized location for the resources to be displayed to users coping with a Title IX case or mental health issue.

Developing the Chatbot:

User-centered design was the focus while developing the ERIN chatbot. In the early steps of the user-centered development, a three-phase model was utilized starting with the “Inspiration Phase” where the designer learns from the user groups they are designing for (What is Human-Centered Design, n.d.; Djamasbi and Strong, 2019). Keeping this phase in mind, proto personas and human-chatbot interaction principles were focuses of the development process, which was assisted through recording and analyzing interviews via Zoom with key informants. The results of the interview sessions are presented within this section, which provides the rationale for design decisions such as the interface, infrastructure, conversational scripts, and proto personas of the target user group. Appendix B provides an illustration of chatbot 1.0.0.

Target Users

Proto personas are “developed through indirect interaction with the user, that is they are assumption based” (Jain et al., 2019). For this project, the proto personas were broken down into cards (sections) of the user’s information using the template provided in “Creating Value with Proto-Research Persona Development” by Jain et al (2019). Given the template and the
assumptions-based model for the proto personas, two key informants were interviewed to develop the proto personas for target users of the chatbot. The first key informant was a project stakeholder for this IQP, who works with college students every day through managing student workers, as well as assisting students with technology problems. The second key informant was also a project stakeholder, working at the Student Development & Counseling Center (SDCC), which handles all therapy cases for mental health, as well as Title IX therapy at WPI. Both key informants provided insight for all the sections for the proto personas: Characteristic Attributes, Background Information, Motivations, Wants, General Frustrations, Preferred Channels of Communication, and Biography. This insight was necessary for focusing the chatbot design on the needs of students that would utilize the chatbot.

After completion of the key informant interviews, three characteristic attributes (i.e. spectrums) emerged that allowed for classification of user groups for this project (Figure 1):

1. Extroversion versus introversion
2. Overachieving versus underachieving
3. Good family experience versus bad family experience
When considering the target users of the chatbot, these attributes provide a view on how the system should function, as well as how their persona will affect their interaction with the system (Jacobs, 2017). Considering the UX of the chatbot, all users have their own personalities making it hard to focus specifically on meeting all needs of each individual user. According to one of the key informants, “WPI is rich with demographics making it hard to meet the needs of all cultures” (Riley-Schafer, 2020). However, by classifying the users into user groups based upon these attributes, designers can focus on meeting the most important needs of user groups (rather than each individual user) making the development effort effective and user focused (Jain et al., 2019).

Additionally, common themes that helped with the development of the chatbot fell into the card categories of motivations, general frustrations (e.g. pain points), and wants of the users. These card categories affected the technological and interaction requirement of the system based
upon user preferences (Jacobs, 2013). Motivations that were present in the assumptions of the key informants were:

1. Seeking help,
2. desire for happiness, and
3. justice.

Many mental health issues and Title IX cases often affect students to the point where functioning could become difficult. And this motivates many students to seek help for a way out or more generally search for a way to feel happy again. Those with Title IX cases may feel as though they are unsafe or wronged; these feelings may motivate them to seek justice. When they finally report an incident, these students may feel good because they feel they stood up against the person that wronged them or made them feel unsafe. In regard to users’ needs (frustrations and wants) that pertains to using the chatbot, the interviews revealed that interacting with the chatbot would be frustrating for target users if they chatbot doesn’t understand the user’s intent, take a long time to respond, and pretends to be a human. What target users would want from a chatbot is to have a way to get immediate help, to be able to submit a request or report on behalf of someone and be able to use the chatbot with little to no instructions (Table 1). The biggest problem with software development is when the user wants and pain points are not identified and addressed by the developer (Rooksby and Ikeya, 2012). Therefore, the user wants and pain points summarized in Table 1, provided an initial roadmap to develop a chatbot that meets the desires of the user and avoids the shortcoming of existing solutions.
Table 1

Assumed General Frustrations and Wants

<table>
<thead>
<tr>
<th>General Frustrations</th>
<th>Wants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Doesn’t understand what is being asked by the user</td>
<td>A way to get immediate help with the chatbot</td>
</tr>
<tr>
<td>Not having a balance between too fast and too slow response time</td>
<td>A way to submit on the behalf of someone</td>
</tr>
<tr>
<td>Not having a balance between human and chatbot</td>
<td>An easy to use application with little instruction</td>
</tr>
</tbody>
</table>

Infrastructure

Designing the infrastructure for the chatbot was an easier task because the chatbot is part of ERIN and as such it is integrated into the ERIN infrastructure. The ERIN infrastructure was designed with scalability and growth in mind; it was designed to be fast, modular, and flexible. The speed was necessary for providing users with the desired results in real time, whereas the modularity and flexibility was necessary to encompass all applications that are available on the ERIN network. For this reason, the design was determined to be a single page application (SPA) to allow for a quick change from one application to another. Node.js as the application framework was utilized to facilitate such a quick change. Furthermore, the Node.js framework provided a modular design by allowing for pre-built libraries to make the development of future applications simpler and hence provided the flexibility and scalability that are needed in applications such as ERIN that are designed to grow based on user needs. Therefore, Node.js technology met the necessary requirements of the infrastructure by having a high performance, being extensible, and having the ability to be scalable.

Chatbot 1.0.0 was designed using the developed proto personas. Two main libraries were utilized for the development of the chatbot: conversational-form and node-nlp. These libraries allowed for the development of the chatbot 1.0.0, which are designed to accommodate the proto
persons. Beginning with the conversational-form allowed for a simple integration of a chatbot user interface (UI) by taking the HTML form and converting it into a text message style conversation (e.g. conversational UI). Furthermore, it provided additional functionality to the HTML form by adding conditionals (e.g. true false statements), as well as the ability to dynamically add responses to the chatbot using the flowCallback function. This provided a foundation for further development of the chatbot to be used in Study 1 to examine user preferences (Figure 2).

**Figure 2**

*Basic default UI of conversational-form library*

Additionally, the use of the node-nlp library allowed for the chatbot to incorporate a NLP functionality to understand the user. This ability to understand was possible due to the library providing the framework to add intents and entities to the chatbot based upon user utterances. The library takes an utterance by the user and compares it to a database of keywords for each
intent available on the chatbot, then returns a confidence score zero to one of the relationships to each intent’s keyword database. Finally, the highest score is taken, and the entity is retrieved by the NLP application. The conversation is then dynamically changed based upon the desired entity of the user. The process can be seen below in Figure 3.

**Figure 3**

*NLP Diagram*

![NLP Diagram](image)

**Interface**

UIs and user experiences cannot be developed perfectly from only one design attempt; they must undergo an iterative design process that improves the design as the project evolves (Djamasbi and Strong, 2019; Nielsen, 1993). The first initial design of the prototype was developed using the proto personas from the key informant interviews. However, this was only a rough prototype until user interviews were completed to verify the personas and get design feedback for the chatbot in Study 1. The proto personas allowed for a few major focuses on the initial development:
1. Develop a welcoming and “calming” design for users that are going through a difficult situation
2. Provide a design that reduces cognitive effort
3. Develop a design that is intuitive and familiar to target users

Based upon the points discussed above, chatbot 1.0.0’s UI was developed using a pastel blue color to provide a welcoming look and feel. To reduce cognitive effort, the conversational UI was surrounded in a border to highlight where the user is supposed to interact with the chatbot. Finally, the conversational UI of the chatbot was created to mimic a text message interface which was a familiar (and hence intuitive) design for the target user group. The final chatbot 1.0.0 design is provided in Appendix B.

**Conversational Scripts**

The fundamental idea of a chatbot is to mimic human conversation on a technology platform that is developed to provide a service (Park et al., 2018). Therefore, the conversational scripts implemented into the chatbot is essential to how well the chatbot operates and they need to be tuned to the service the chatbot is utilized for. Based upon previous research, it is essential to include empathy into a conversational UI to achieve humanistic conversation (Lin et al., 2020). Additionally, due to the nature of the service that the chatbot in this IQP is being developed for, it is crucial to incorporate empathy into the chatbot. The conversational scripts for the ERIN chatbot were developed using training manuals for WPI’s SDCC Student Support Network program, as well as the research paper, “What is Empathy?” by Beatrice Kalsich, with examples of empathetic responses. Once the scripts were developed, they were reviewed and discussed during another key informant interview with an expert therapist at the SDCC. Using the notes from this interview, the scripts were refined to remove any human feelings that the
chatbot couldn’t feel to balance the humanistic attributes of the human-chatbot interaction.

Appendix C presents the finalized conversational scripts that were implemented into the chatbot.

The next step was to conduct user studies to refine the design of the chatbot. The following sections discuss the user studies that were conducted throughout this IQP.

**User Studies:**

**Study 1 – User Interviews:**

*Overview and Design*

This user study was conducted to 1) verify or refine the proto personas that were developed via key informant interviews, 2) understand users’ knowledge of Title IX, and 3) obtain information about user design preferences (e.g. color scheme).

Research personas are developed by conducting user research such as one-on-one interviews with users (Jain et al., 2019). This approach was used to verify/refine proto personas. As in prior research (Jain et al., 2019), proto personas were used to design interview questions to get relevant information such as family life, involvements on campus, and stressors during the day. Using these questions, which directed the conversation in an open-ended way, helped to obtain the desired information needed from the user to verify the proto personas.

Since the chatbot provided resources related to Title IX, a government defined law, it was important to examine whether target users were familiar with Title IX. Therefore, the second section of the user interviews had interview questions to obtain background knowledge the users have for Title IX. For example, participants were asked whether they knew what Title IX is.

Formative design process is the idea that you ask a user to identify any improvements, weakness, or details for the development of a design through prior experiences and/or rough drafts of current design prototypes (Tullis and Albert, 2013; Reigeluth and Frick, 1999). Thus,
chatbot 1.0.0 was presented to the user to get user feedback. Furthermore, there were questions (e.g. Looking at the initial design, how would you alter the design?) to gain insight on their previous experiences with chatbot to determine what to avoid in the next iterative design.

Understanding the user is an essential part for the chatbot to determine the intent of the user and to give the user the correct entity. Therefore, the user was presented with four scenarios which were developed based upon the key informant interviews. Each participant was asked to read the scenario and think out loud about what they would expect to say to the chatbot and the responses they expect to get back from the chatbot.

**Participants**

Key informants suggested that the typical user of the ERIN chatbot would be college students of the age between 17-25 years old. For this study, users within this age range were recruited by reaching out to them through professors teaching classes and by reaching out to clubs at WPI. All users were contacted via email asking them to participate in an online user interview. The study incorporated 15 college students with varying backgrounds and characteristics that met the desired user groups (undergraduate students with an age range between 17 to 25). The collected information from these interviews was used to verify or refine proto personas and/or to create new research personas. The interviews were also used to collect user utterances for the NLP keyword databases, and to obtain design preferences for developing the chatbot.

**Measurement**

Consistent with typical research techniques for interviews (Wargo, 2013), the process of coding the interviews were utilized for analyzing the results of the interviews. The coding process was completed by creating transcripts of the user interviews and looking for natural
meaning units given by the participants. These natural meaning units refer to “self-delimiting expressions in the words of participants” (Wargo, 2013). Using these natural meaning units, common themes were identified between participants and recorded by the investigators. Lastly, the user utterances were recorded and broken down into keywords that were incorporated into the databases to get the desired entity out of the chatbot.

**Results**

The interview process was conducted by asking open-ended questions about their daily life. The analysis of user responses verified the proto personas that were developed through the interviews with the key informants, no new personas emerged from this analysis. The interview questions also captured design preferences. All transcripts were coded to find common themes within the “wants”, “general frustrations”, “background knowledge/experience”, and “color preferences.” These common themes were utilized in the further development of the chatbot to make a user-centered design.

The “wants” of users are particularly important to UX design because they provide a fundamental view into their perspective towards a problem (What is Human-Centered Design?, n.d.). The majority of participants stated that they would prefer the chatbot to be anonymous (n=11), and easy to use (n=11). A good portion of participants wanted the chatbot to incorporate a tutorial for instructions (n=8). Less than half of participants (n=7) stated that they would like to have a way to see the chat history through scrolling and have the ability to talk to a human agent when the chatbot gets stuck. Some participants (n=6) wanted the chatbot to be empathetic and personable, not pretending to be human (balance of human and chatbot), understand user’s intent, and have a text messaging s interface. The results are displayed below in Figure 4.
Learning from past experiences allow for products to adapt and evolve in positive ways. These experiences present pain points towards a product that users wouldn’t like to see again. During the user interviews, users were asked to explain their pain points based upon previous use with a chatbot and/or their interaction with WPI’s website. With these in mind, the majority of participants disliked when developers use a small font and bright colors, specifically red (n=11). Additionally, five students didn’t like when the wording of the conversation was odd or redundant compared to their normal language. Furthermore, there were three students who stated they felt confused and overwhelmed when they interacted with a design that had distracting features or when they received unprompted information. Figure 5 provides the results of these common pain points.
As mentioned earlier it was important to understand the level of students’ familiarity with Title IX because the chatbot provides information about this topic. Thus, participants were asked whether they knew what Title IX was. The majority of users (n=13) were aware of Title IX, with an 86.67% of users not only knowing what Title IX was but also being trained how to handle Title IX situations (Figure 6).

Similarly, it was important to tap into users’ prior experience with chatbot. Interestingly enough, our results showed that only about 60% of participants (n=9) had prior chatbot experience. This indicated that it may be a good portion of target users first time using the chatbot (Figure 6).
The interviews with key informants indicate that it was necessary to provide a welcoming feel to prevent heightening emotions with design choices. Therefore, the users were asked to give their perceived view on colors to determine what was the most welcoming. Almost all the users (n=14) felt that blue was a welcoming color; however, they didn’t want anything darker. They preferred lighter pastel colors when they described welcoming colors. Additionally, green and yellow were sought out to be welcoming colors (n=8 and n=7 respectively) presenting an idea of happiness. This trend is presented within the bar graph of Figure 7 below.
Study 2 – User Experience and Usability Testing:

Overview

After completion of Study 1, the user-centered focus was shifted towards the second phase known as the “Ideation Phase,” which the idea of taken the research personas (Appendix D), design preferences, and prior experience to develop possible solutions (What is Human-Centered Design?, n.d.; Dajasibi and Strong, 2019). Using the results from Study 1, the design was altered to chatbot 1.0.1 by incorporating common wants, utilizing suggestions, and avoiding pain points. The first initial changes were made to the user interface, making the chatbot’s conversation UI smaller with the border on both sides. Additionally, the avatars were slightly modified by increasing their size (making them more prominent) and removing the thumb (e.g. avatar background circle to create a simpler design) for ERIN to reduce cognitive effort of the users. Then, regarding the text message style design determined in Study 1, the font size was
changed to a san serif and increased in size to reduce cognitive effort. The area for the chatbot’s conversational UI was decreased to create a more pleasing design. Lastly, the background color of the chatbot’s conversational UI was changed to white to improve the contrast and readability, which again reduces the cognitive effort of the user. Furthermore, functionality was updated by adding more keywords to the database by using the user utterances that were collected. Appendix B provides an illustration of chatbot 1.0.1.

Given the new version of the chatbot, UX and usability testing was the next step for the ERIN chatbot. User experience and usability are important to measure because they can reveal how effective the chatbot is at reaching the user’s goal and how design improvements can increase technology adoption (Djamasbi and Strong, 2019). User experience of a technology can be captured using surveys such as system usability scale (SUS) and mobile user experience (MUX). Similarly, users’ perception of and reaction to the usability of a design can be captured by variables such as behavioral intention, task effort, usefulness, and ease of use (Komninos, 2020; Jain et al., 2020). Additionally, the chatbot was developed to be utilized on different technology mediums, which was integrated into the usability testing procedure to create a between user study design. In order to effectively measure the usability of the application, the usability testing procedure included both qualitative and quantitative methods, which are described in the following sections.

**Participants and Design**

Like Study 1 participants of the user interviews, users were recruited in a similar process and contacted via email. The study incorporated 12 undergraduate college students using a between study design. Students were randomly assigned to one of the two treatments: Six students were randomly assigned to test the chatbot on a laptop and six students were randomly
persons assigned to test the chatbot on a mobile phone. All participants were asked to complete one task (use the chatbot to find information) based upon a given scenario.

Understanding what the chatbot can do for a user is an influential part of the experience with the chatbot. Users’ should be able to tell the functionality of the chatbot just by their initial impressions. Users’ first reaction to the chatbot was measured qualitatively using the problem discovery method, in which users are asked to view a technology artifact and report (or think out loud) what they think they could do with that technology. In this part of the study, users’ verbal and non-verbal reactions were recorded by the experimenter and coded for analysis.

User experience (UX) is “the subjective experience of an individual in his/her encounters with a technology [medium]” (Djamasbi et al., 2016). Both UX and usability heavily depends on the context (Djamasbi and Strong, 2019). Therefore, all users were presented with a scenario (context) and asked to use the chatbot given that prompt. After completing the scenario with the chatbot, two UX surveys (Tullis and Albert, 2013; Wilson and Djamasbi, 2019) were used to capture users’ experience of using the chatbot.

Usability in this IQP focused on cognitive effort and users’ willingness to use the chatbot. Therefore, this IQP used surveys that can measure task effort (finding resources via chatbot) (Wang and Benbasat, 2009) and technology acceptance behavior (Komninos, 2020; Jain et al., 2020). All these measurements are explained in more details in the next section of this paper.

Measurement

Problem discovery (whether users could guess what the chatbot can do for them) and resource quality (whether the chatbot provided correct information based on user input) was captured by the experimenter through observation. The same coding process as in Study 1 was utilized to determine whether the user knew what the application’s functionality was upon initial
impressions and if they received the desired resources from the chatbot based the conversation that they had with the chatbot.

Self-reported measures were used to capture task effort (Wang and Benbasat, 2009), technology acceptance (Hang and Yan Tam, 2006; Adipat et. al, 2011), and user experience (Tullis and Albert, 2013; Wilson and Djamalsbi, 2019). Users were asked to rate their responses on a 5-point Likert with one correlating to strongly disagree and five correlating to strongly agree. These quantitative measures are listed below:

1. Perceived Task Effort (PTE) (Wang and Benbasat, 2009)
2. Perceived Ease of Use (PEOU) (Adipat et al. 2011)
3. Perceived Usefulness (PU) (Adipat et al. 2011)
5. System Usability Scale (SUS) (Tullis and Albert, 2013)
6. Perceived Behavioral Intention (PBI) (Hang and Yan Tam, 2006)

Results

Each of the surveys used in this IQP consisted of several item. Consistent with UX literature in information systems (IS), average values of items for each survey was created for each participant. For SUS, this average value was called ASUS. This is because SUS items are typically translated to SUS scores which can range from 0 to 100 (Tullis and Albert 2013). Sus scores were also calculated. Additionally, for MUX, a single score like SUS was calculated (MUXS).

T-Tests were used to analyze the impact of the technology medium (laptop versus mobile phone) on usability of and experience with the chatbot. The results showed that PBI was rated
almost significantly \((p=0.06)\) better in mobile phone group compared to the laptop group.

Similarly, the results showed that SUS score was rated almost significantly \((p=0.07)\) which is again better in the mobile phone group compared to the laptop group. While the mean of PEOU, PU, PTE were not significantly different in the two groups, the analysis shows better ratings in the mobile condition. ASUS vs. SUS and MUX vs. MUXS had identical results. It is important to note that the SUS score for laptop was in good range \((79.58)\) and for mobile phone was in excellent range \((91.67)\) indicating that both mediums provided positive user experience (Bangor et al., 2009). These T-Test results are provided in Table 2.

Table 2

*T-test data analysis of usability testing (assuming equal variances)*

<table>
<thead>
<tr>
<th>Survey Title</th>
<th>Medium of Technology</th>
<th>Mean</th>
<th>Std Dev</th>
<th>df</th>
<th>t Stat</th>
<th>P(T&lt;=t) two-tail</th>
</tr>
</thead>
<tbody>
<tr>
<td>PTE</td>
<td>Laptop</td>
<td>4.28</td>
<td>0.39</td>
<td>10</td>
<td>0.43</td>
<td>0.67</td>
</tr>
<tr>
<td></td>
<td>Mobile Phone</td>
<td>4.39</td>
<td>0.49</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PEOU</td>
<td>Laptop</td>
<td>3.83</td>
<td>0.72</td>
<td>10</td>
<td>1.75</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td>Mobile Phone</td>
<td>4.44</td>
<td>0.46</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PU</td>
<td>Laptop</td>
<td>4.28</td>
<td>0.71</td>
<td>10</td>
<td>1.66</td>
<td>0.13</td>
</tr>
<tr>
<td></td>
<td>Mobile Phone</td>
<td>4.83</td>
<td>0.41</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MUXS</td>
<td>Laptop</td>
<td>83.33</td>
<td>8.74</td>
<td>10</td>
<td>0.29</td>
<td>0.78</td>
</tr>
<tr>
<td></td>
<td>Mobile Phone</td>
<td>85.42</td>
<td>15.08</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SUS</td>
<td>Laptop</td>
<td>79.58</td>
<td>10.54</td>
<td>10</td>
<td>1.99</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td>Mobile Phone</td>
<td>91.67</td>
<td>10.45</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MUX</td>
<td>Laptop</td>
<td>4.33</td>
<td>0.35</td>
<td>10</td>
<td>0.29</td>
<td>0.78</td>
</tr>
<tr>
<td></td>
<td>Mobile Phone</td>
<td>4.42</td>
<td>0.60</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASUS</td>
<td>Laptop</td>
<td>4.18</td>
<td>0.42</td>
<td>10</td>
<td>1.99</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td>Mobile Phone</td>
<td>4.67</td>
<td>0.42</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PBI</td>
<td>Laptop</td>
<td>2.67</td>
<td>1.23</td>
<td>10</td>
<td>2.14</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>Mobile Phone</td>
<td>4</td>
<td>0.89</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Next, technology acceptance model was examined via regression analysis (Table 3).

TAM asserts that PBI is predicted by PU and PEOU. The results show that PBI was almost significantly (p=0.06) predicted by PU but not by PEOU. PU, however, was significantly predicted by PEOU.

Table 3

*Regression data analysis*

<table>
<thead>
<tr>
<th>Model</th>
<th>Adjust R Square</th>
<th>Unstandardized Coefficients</th>
<th>Standard Error</th>
<th>t Stat</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( PBI = B_0 + PEOU \times B_1 )</td>
<td>0.17</td>
<td>0.93</td>
<td>0.52</td>
<td>1.80</td>
<td>0.10</td>
</tr>
<tr>
<td>( PBI = B_0 + PU \times B_1 )</td>
<td>0.25</td>
<td>1.11</td>
<td>0.52</td>
<td>2.14</td>
<td>0.06</td>
</tr>
<tr>
<td>( PU = B_0 + PEOU \times B_1 )</td>
<td>0.79</td>
<td>0.85</td>
<td>0.13</td>
<td>6.53</td>
<td>0.00006</td>
</tr>
</tbody>
</table>

Next, the impact of PTE on each of the three TAM variables were examined separately, i.e., via three linear regression models (Table 4). The results do not show that PTE has a direct impact on PBI (p=0.10) but it has an almost significant impact on PEOU (p=0.07) and a significant impact on PEOU (p=0.04).

Table 4

*Regression data analysis*

<table>
<thead>
<tr>
<th>Model</th>
<th>Adjust R Square</th>
<th>Unstandardized Coefficients</th>
<th>Standard Error</th>
<th>t Stat</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( PEOU = B_0 + PTE \times B_1 )</td>
<td>0.22</td>
<td>0.83</td>
<td>0.41</td>
<td>2.03</td>
<td>0.07</td>
</tr>
<tr>
<td>( PU = B_0 + PTE \times B_1 )</td>
<td>0.31</td>
<td>0.88</td>
<td>0.37</td>
<td>2.41</td>
<td>0.04</td>
</tr>
<tr>
<td>( PBI = B_0 + PTE \times B_1 )</td>
<td>0.17</td>
<td>1.44</td>
<td>0.80</td>
<td>1.81</td>
<td>0.10</td>
</tr>
</tbody>
</table>
Finally, throughout the usability testing all participants (n=15) were able to determine what the chatbot was used for upon entering the landing page of the chatbot. Furthermore, the chatbot had a success rate of 100% by given the desired resources to the user every time (resource quality).

**Discussion**

Data analysis showed that while there were no significant differences in task effort, ease of use, and usefulness, participants in the mobile treatment provided more favorable ratings. Participants’ SUS score was in a good range for the laptop group and in an excellent range for the mobile groups; the differences in the SUS score were almost significant (p=0.07). The differences in intention to use the chatbot were also almost significant (p=0.06) between the groups with people in the mobile group indicating that they were more likely to use the chatbot in the future. Today’s society emphasizes using a mobile phone for many different tasks from communication to entertainment (Mobile Phones and Society, 2016). This trend in society maybe a reason for better PBI ratings in the mobile phone group. Additionally, chatbot was designed to have a text message style interface common in mobile phones. This may be a reason for better SUS ratings in the mobile phone group.

The results showed that behavior intention was predicted (almost significantly) by usefulness but not ease of use. The usefulness of the chatbot, however, was heavily influenced by its ease of use. In other words, cognitive effort (ease of use) of the design had a significant impact on how useful the chatbot was perceived by the participants. While the PEOU ratings in both laptop and mobile groups were in the above average range (<3.33), ratings in the mobile groups were more positive in the mobile group (4.44 out of 5) than in the laptop group (3.83 out of 5). Usefulness ratings in both groups were more similar (4.28 and 4.83) indicating highly
positive perceptions. These results suggest that regardless of the medium the chatbot was perceived as highly useful. The average PBI ratings in the laptop group (2.67 out of 5) was in the average range (2.33–3.6) while in the mobile group (4 out of 5) was in the above average range (<3.67); the differences in these ratings were almost significant (p=0.07).

The results show a better user experience and usability ratings in the mobile group indicating that the chatbot maybe more frequently utilized by the target users when accessed via mobile phones. This aligns with previous research that states, “usability affects the satisfaction which in turn affects intention to use” (Belanche et al., 2012). Together, the results suggest that the medium (laptop versus mobile phone) may have an impact on how successfully the chatbot is adopted.

TAM, which has been an extensively tested, is an influential model for assessing user acceptance of a product (Charness and Boot, 2016). Our results showing a significant relationship between PEOU and PU is consistent with TAM. However, our results showed that the relationship between PBI and PU was only almost significant (P = 0.06). One possible explanation is that the sample size was low in this study (total of 12 users, six users per technology medium). These results have the potential to become significant with more users completing the usability testing. Another possibility is that the acceptability of sensitive applications, such as the one developed in this IQP, may depend on variables that were not captured in this project.

The non-significant relationship between PBI and PEOU in this project may be due to the task. Research shows that the task presented to the user could sway the data for the PBI and PEOU relationship making it not significant even with a larger sample size (Dishaw and Strong, 1999).
Limitations

Testing was conducted using a small sample size of users. To verify the chatbot further, the user tests would need to be conducted with a much larger sample size of students (e.g. at least 15 users per medium of technology). However, the results obtained from a small sample size in this project are encouraging.

The results of this IQP shows that while task effort can have a direct impact on ease of use and usefulness, it does not impact behavioral intention directly. Future studies are needed to examine the impact of task effort on TAM more thoroughly.

User tests were conducted during the COVID-19 pandemic. The presence of the pandemic could have affected the results. Future studies when the COVID-19 pandemic is concluded are needed to test this possibility.

Chatbots are typically conversational in nature, which makes it seem more of a user conversation. The conversation for this project was limited to finding resources. Future research is needed to examine the suitability of the platform for providing support in addition to providing broader resources related to Title IX and mental health related issues. Additionally, the conversational aspect of the chatbot needs more design and testing iterations to improve the adoption of the chatbot.

Future Work

Since the chatbot was developed using libraries for Node.js, it presents a flexible and extensible design that makes growth possible. The ERIN chatbot will be able to change with the resources that are present in the WPI community. Additionally, due to its scalable design, the ERIN chatbot will be able to expand the resources and natural language processing to support finding resources for other services on the campus such as academic, food courts, etc. This
essentially means that the chatbot could potentially serve as the main hub for all resources at WPI allowing the WPI community to obtain desired resources in a quick, conversational, and effective way. The chatbot can also be translated into support in other organizations.

The data and user comments obtained from the usability testing will be utilized in the next version of the chatbot. The chatbot development will follow the user-centered design by continuing through an iterative design and test process that meets the needs of users and is appealing to users.

The NLP database can be altered to incorporate machine learning making it more effective at obtaining the desired resources of the user. By recording the utterance of the user and the resources they choose, the chatbot can continually update its database to tune the databases towards the desired entities based upon the direct intents recorded. This would allow the confidence score to be more accurate and prevent false answers that the users are not looking for.

Additionally, future studies will be conducted using the information obtained from this project to verify the trends that were observed throughout the duration of this IQP.

Contribution

Theoretical

Communication through technology is often assumed to lack social presence (Calefato and Lanubile, 2010). The chatbot can improve social presence because it utilizes a conversational UI that the user interacts with. This idea is supported by studies that show many users preferred communicating with a chatbot rather than going through a webpage because users can talk to a chatbot as if they are talking to a friend (Brandtzaeg and Folstad, 2017). The
results of this IQP suggest that medium (laptop versus mobile phone) may also have an impact on perception of social presence.

**Practical**

Using the results from the usability testing, the SDCC website could be reorganized to reduce users’ cognitive load. Currently, the SDCC website is a subpage on the main WPI website with many different subpages without explicit way to access the resources on the subpages. Based upon the results of the usability testing in this project, SDCC webpage may benefit from a phone application that provides SDCC resources explicitly labeled with an emergency contact as well.

The chatbot framework could be utilized in other domains. For example, the chatbot can help in emergency response situations such as COVID-19 outbreak. Because the chatbot framework makes it easy for users to ask questions (e.g. should I take a test when I don’t have any symptoms?), it can help people find resources in a more empathetic way than using a website (e.g. looking through FAQ). The conversational interface of the chatbot may also serve as a great tool for individuals with disabilities. Because people can interact with the chatbot in a familiar conversational way, the chatbot is likely to create a low cognitive effort and empathetic way for people with disabilities to find resources they need.

Overall, the chatbot could be utilized for many different services and applications in a natural way rather than just having a neutral webpage. Chatbot can provide an interactive experience in a timely and helpful fashion. This IQP was just one example of how a chatbot can be incorporated into a service; however, not all technological problems can be solved using a
chatbot and should be considered when deciding upon what technology platform is best for a specific situation.
Bibliography:


https://www2.ed.gov/about/offices/list/ocr/docs/investigations/open-investigations/tix.html


https://www.researchgate.net/publication/330741762_Formative_Research_A_Methodology_for_Creating_and_Improving_Design_Theories


Wang, Weiquan, and Benbasat, Izak (2009), "Interactive decision aids for consumer decision making in e-commerce: the influence of perceived strategy restrictiveness", MIS Quarterly, 33, 293-320.

What is Human-Centered Design? (n.d.). Retrieved May 21, 2020, from Design Kit:
https://www.designkit.org/human-centered-design

Appendices:

Appendix A – ERIN (erin.wpi.edu)
Appendix B– Design Iterations

Chatbot 1.0.0

Chatbot 1.0.1
Appendix C – Conversational Scripts

Greeting/Refining Question 1:
Hey there! My name is ERIN, your friendly robot assistant! Could you please explain what happened?

Refining Question 2:
It takes a lot of strength to persevere through these sorts of things. I want to help you through this, please choose a category so that I may assist you further!

Gender Discrimination
Situations like that can be frustrating, I am on your side. Here are some resources to help overcome this.

Sexual Discrimination
People can act in inconsiderate ways and they shouldn’t act that way. I would like to help you obtain help through these resources.

Racial Discrimination
That kind of treatment is unacceptable. Let’s get you some support to help stand against this!

Nationality Discrimination
It’s totally natural that you would feel this way. Here are some resources to help you with these feelings!

Age Discrimination
It is understandable to feel hurt about this. You shouldn’t have to go through this alone, here are some resources to assist you!

Retaliation
The actions taken towards you were not warranted and it must be frustrating to go through. I want to help you with this!

Anxiety
I know it feels like no one understands, but I am here for you and I want to give you resources to help calm these thoughts and feelings.

Depression
That’s definitely a challenging set of circumstances. Based on what you’ve described, I wonder if these resources might help.

General Stress
It has to be difficult to deal with this right now. It’s normal to struggle when it feels like there’s too much going on, here’s some resources that can help when that’s the case.

Gambling
It takes a lot to admit this and look for help, I have found some resources that could help you get to the next step!

Substance Abuse/Addiction
I know it isn’t easy to deal with stuff like this, but You are making a tremendous first step by talking to me about this issue! I have found some resources to help you continue to a healthier lifestyle!

Reporting for a Friend/Worried about a Friend
It takes a special friend to want to help. I have found some resources to help you achieve this help!

Eating Disorder
It is normal to have concern with this situation. I have found some resources to help alleviate some of these concerns!

**Sleeping Disorder**
Dealing with the effects of this can be difficult. I want to help you with this, and I have found some resources that could be of assistance!

**Sex Violence/Harassment**
This is a difficult situation and I think you’ve shown a lot of courage and strength in how you’re handling things. I’m here for you and I have gathered some potential resources to help you!

**Self-Harm/Suicidal Thoughts**
It’s way more common than people think to struggle with these kinds of thoughts, and it’s important to seek help in dealing with them. Here are some resources to help!

**Closing Remark:**
Thank you for allowing me to help you take these initial steps in the service of your wellness. I hope it is the start of some positive change for you! It means a lot to me! Have a wonderful day! I am always here to help!
Appendix D – Personas

Note: * indicates an assumed card type due to the limitations of user interviews

### Emma Smith

**User Concern:** Stress

**Happiness**

**Motivations**

- Seeking Help
- Justice

**Wants**

- She would like an easy to use chatbot that is simple.
- She hasn’t used a chatbot before and would like a tutorial to explain how to use it.
- She would like the chatbot to be quick since she doesn’t have much time to spend on the chatbot.

**General Frustrations**

- Doesn’t like when the chatbot is not easy to read due to size.
- Doesn’t want a red design because it comes off as aggressive.
- When a chatbot leads them in the wrong direction.

**Preferred Channels of Communication**

- Phone
- Social Media
- Email

**Goal**

Emma is involved in many different extracurricular activities, as well as college workload so she has little to no time. Due to the activities she is doing, she has developed an overwhelming stress trying to balance school and activities. She is looking for a way to help her balance her time and alleviate some of this stress that has been occurring throughout the semester.

### Robby Jones

**User Concern:** Sexual Discrimination

**Happiness**

**Motivations**

- Seeking Help
- Justice

**Wants**

- Wants to be able to choose between multiple resources rather than one.
- Wants to be able to submit a report anonymously so that they don’t get backlash from others.
- Wants a chatbot that is nice and kind towards them while conversing.

**General Frustrations**

- Dislikes when you must submit a specific command rather than any text inputted.
- Doesn’t like distracting designs.
- Doesn’t like rude conversations with machines.

**Preferred Channels of Communication**

- Phone
- Social Media
- Email

**Goal**

Robby is a LGBTQ advocate and his professor discriminated against him saying derogatory terms about his sexual preferences. They want to be able to submit a report without being placed to the case because they are scared of the backlash they may receive from the student population or family members.
Adam Flattery

User Concern: Depression

General Frustrations:
- Doesn’t like when a chatbot give unprompted information
- Doesn’t like bright colors on a design
- Doesn’t want there to be redundant or odd wording

Age: 19 years old
Occupation: Full-time Student
Location: Massachusetts
Characteristics: Sigma Pi Brother, Member of Alpha Beta Honor Society, President of Symphony Association, Part of String Ensemble, Introvert, Overachiever, Family Doesn’t Get Along Well

Goals:
- Adam is a Sigma Pi brother and has been struggling with depression with the weight of college classes, problems with his family situations, and other involvements on campus. He is needing a way to help him get those emotions out of his thoughts through appointments.

Happiness: 4/5
Seeking Help: 3/5
Justice: 3/5

Motivations:

Wants:
- Would like to have a chatbot that mimics a text message, so he is comfortable with using it.
- Wants to be able to talk to a user agent if something doesn’t work correctly for him.
- Wants to have a balance between humans and robot.

Christina Johnson

User Concern: Sexual Harassment

General Frustrations:
- Doesn’t like when the chatbot is very defined by what is coded and not flexible
- Doesn’t like the idea of using images with a chatbot because it can be offensive
- Doesn’t want the chatbot to be a barrier between what she is trying to do

Age: 20 years old
Occupation: Full-time Student
Location: Oklahoma
Characteristics: CA, Alpine Ski Team, Hiking, Camping, Overachieving, Introvert, Divorced Parents

Goals:
- Christina was a victim of a sexual harassment case. There was a hidden camera found in her room and she is really worried about what may have been caught on the camera, as well as who may have been watching her. She wants to get justice and closure on the situation but is unsure about how to go about reporting the case to the university.

Phone
Social Media
Email

Preferred Channels of Communication

Happiness: 3/5
Seeking Help: 4/5
Justice: 3/5

Motivations:

Wants:
- Wants to be able to see the chat history
- Wants to have some sort of customization for the users
- Wants to have everything on the chatbot if it can be