Tampere University

Li-Chia Cheng

# AN EXPLORATORY DESIGN OF STRESS RELIEF CHATBOTS FOR TAIWANESE ADOLESCENTS

Faculty of Information Technology and Communication Sciences
Master's Thesis
June 2021

# ABSTRACT

Li-Chia Cheng: An Exploratory Design of Stress Relief Chatbots for Taiwanese Adolescents
Master's Thesis
Tampere University
Human-Technology Interaction
June 2021

---

Currently, there is a growing interest in the application of chatbots for mental health support, both for adults as well as children. Multiple studies have been done in regards to the domain of chatbot for mental support for adolescent youths, but there is a dearth of knowledge in regards to culture-specific needs as most related studies have been done with a western audience in mind. This study, therefore, tries to examine what are the particular needs of Taiwanese adolescents and how to go about designing a stress support chatbot, by the way of an exploratory design study. Taiwanese middle school counselors were interviewed to provide insights for the development of a prototype chatbot that utilized Google Dialogflow's natural language processing to identify stressors based on keywords and give support to the Taiwanese adolescent target group. The analysis of both qualitative and quantitative results collected from user testing revealed that due to the power distance that is prevalent within Taiwanese society, a chatbot's persona determines the role it plays and the needs it can fulfill for Taiwanese adolescents looking for stress support. Quantitative findings further detail what chatbot interactions are desired user experience in the context of stress relief for Taiwanese adolescents.

Keywords: Chatbots, Conversational agents, Cross-Cultural Design, Conversational design, Stress

**CONTENTS**

## 1. Introduction

Mental illness in the form of depression and anxiety is common among the younger population. According to the WHO, mental illness accounts for up to 16% of the disease and injury in 10 to 19-year-olds worldwide (WHO, 2020). Globally, half of all mental conditions start at the early age of 14, but most are untreated and unreported (Kessler et al., 2007). A survey conducted in 2013 showed that American teens have reported feeling more stressed than is healthy and the trend seems to have continued, with similar reports in 2018 and 2020 (APA, 2013; APA, 2018; APA, 2020). While arguments as to whether Asian or western adolescent populations experience higher stress levels have been made for both sides (Crystal et al., 1994), differences in both causation and coping methods have been reported (Zgambo et al., 2012). Due to cultural differences, ethnic Chinese adolescents tend to experience conflicts regarding independence in which traditional values that stress obedience and collectivism clash with a growing individualism that stems from the growth of western values. (Hsieh et al., 2014). Compounding the issue further is the presence of power distance, where social hierarchy dictates at times the way people interact with one another. Taiwan, where 95 to 97 percent of the population is ethnic Chinese, experiences that same culture clash, resulting in complexities in mental illness treatment. Further exacerbating the issue is the perception of counseling, which is viewed as a last resort for those only with severe issues or lacking in personal network support by young adults in Taiwan, leading to continued stigmatization of mental issues and help-seeking (Lin, 2001). Rarely have the Taiwanese population sought out counseling on their own and it is largely through the use of mandatory counselling when their home teachers request troubled students to go to the school's counselor's office do Taiwanese adolescents experience professional mental support.

One way of tackling the unique issues caused by the cultural background in Taiwanese adolescents could be the use of conversational agents. Conversational agents are software systems with which users interact with via either text input or spoken input. The system then responds to the user by mimicking human language interactions to the user's input. As the interactions mimic traditional face-to-face therapy interventions, chatbots could be effective in providing mental health support (Fitzpatrick, 2017). An additional upside of conversational agents is the fact that they provide availability through multiple touchpoints such as social media, messaging applications as well as dedicated mental health applications (mHealth).

Prior studies have been done on the design and development of chatbots for mental health support for adolescents such as Woebot (Fitzpatrick, 2017) and Teenchat (Huang, 2015), as well as for adults (Park, 2018; Inkster et al., 2018). However, there are two knowledge gaps from prior studies. The first gap of knowledge is as most studies done in the domain of chatbots for mental health are primarily focused on the implementation and outcome effectiveness, there are relatively few studies on what principles of the conversation designs are used for mental health support. The second gap is that most recent studies are mainly conducted in western countries, which may lead to a biased approach in terms of the effectiveness of the chatbot in a multicultural approach. As stated prior, culture can affect both stress causation as well as coping methods which means that further knowledge is needed to appeal to and provide relevant support for targeted users of different cultural backgrounds. As the conversation characteristics of chatbots deal with the use of language as an interaction, conversation design focus for specific users needs to be kept in mind, both in the domain of mental healthcare as well as the culture of the users, which in this study is Taiwanese adolescents.

This study, therefore, seeks to extend the research on the design of chatbots for stress relief and intervention, specifically the adolescent population in Taiwan. Through the design and testing of a prototype conversational chatbot, the study will explore how chatbots can assist Taiwanese adolescents in stress relief and what needs should the chatbots fulfill. In particular, this thesis will aim to answer the following research questions:

- What roles should a chatbot for stress relief support for adolescents in Taiwan fulfill to address their specific needs?
- What makes a good user experience for Taiwanese adolescents using a text-based chatbot for stress relief?

To answer the research questions above, this thesis is structured as noted: Chapter 2 Related Work will help set the theoretical foundation of the study through reviews of academic and industry literature that provides the background of adolescent stress viewed from the Taiwanese cultural lens as well as the history of chatbots in general and their implementation in the mental health sector. Chapter 3 presents the methodology the thesis will use to answer the research questions, including the approach of design as well as evaluation methods. Chapter 4 will discuss the initial research results synthesized from interviews with local mental health professionals. Chapter 5 presents details on the design and implementation of the prototype chatbot. Chapter 6 presents and analyzes the user test evaluation results of the prototype chatbot. Chapter 7 synthesizes the overall findings and presents them in discussion. Chapter 8 concludes the research, and also presents limitations as well as suggests future research surrounding chatbot design.

## 2. Related Work

This section details the literature works and background of the study beginning with how culture affects expression of stress and how stress is perceived by the general populace in Taiwan. The second part of this section discusses chatbot history and current studies done with regards to its use within the domain of mental health.

### 2.1 Stress and Culture

Stress is a part of our everyday lives. It occurs when we experience a change in situations and thus requires the body to adjust or respond to said change. Stress helps our body stay alert, keeps us motivated, and focuses our mind to avoid danger. Holistically, we can view stress as a necessary part of our lives. However, continuous stress without relief becomes dangerous, leading to the feeling of being overwhelmed and can trigger increased emotional distress (depression, anxiety), antisocial behaviors (delinquency), risky health behaviors (substance abuse, sexual risk-taking), and even detrimental physical health (skin rashes, lower immunity to illness). A part of what makes mental disorders so prevalent in adolescents is due to increased stressors they experience during the time of growth both mentally and physically as well as being underdeveloped in terms of coping mechanisms.

While studies done in prior decades have brought about a better understanding of how adolescents cope with stress, most of these studies continue to be based on individuals with western backgrounds, leading to a bias in the knowledge of how to approach adolescent stress. Only recently has there been a bigger awareness of how culture plays into the context of human behavior. Based on Hofstede's cultural dimensions, Matsumoto (1989) theorized how cultures affect the perceptions of emotions. Cultures that are scored high on individualism (expression of self), low power distance (hierarchy of social positions), and low uncertainty avoidance (tolerance for uncertainty and ambiguity) can be described as individualistic cultures. On the other hand, cultures that were the opposite can be categorized

as collectivist cultures.

Western culture is more individualistic in nature, with one's self-identity being unique and independent from others through the focus on one's personal needs and desires. Expressions of self and of emotions may even be encouraged as the expression of emotions is attributed as unique individualism. Eastern culture on the other hand is more collective. The value of an individual is not so much on oneself, but rather how one relates to and works in tandem with others in a bigger picture. Fitting in and harmonious living with your society, therefore, is not only valued but rather expected of the individual in question. Collectivistic cultures frown upon the communication of negative emotions as these emotions are perceived as a threat to the existing social order. This paradigm shift of thinking regarding human behavior in conjunction with the difference in culture therefore also affects how we should tackle the issue of adolescent stress since a large part of adolescent stressors is due to the development of self-identity and expression of emotion that comes with it.

**2.2 Taiwanese Youth Stress Factors and Perception**

Coming from a collective culture, the adolescents in Taiwan experience this to a certain degree. Like many adolescents globally, the youths of Taiwan are prone to mental disorders from the increase of stress during their period of growth. The cause of this is a combination of academic performance pressure, interpersonal relationships, and familial relationships (Wang, 2007). The top five self-reported stressors were: 1. too many tests, 2. comparisons to others from parents, 3. uncertainty of the future, 4. not enough sleep, and 5. not enough time. Wang's study (2007) surveyed 1278 middle school adolescents, concluding that for Taiwanese middle schoolers, stress relief and resilience depended largely on the individual's social support and only minorly dependent on self-improvement and self-control.

In this sense, the mental wellbeing and stress resilience of Taiwan's adolescents correlates to how well the person fits within social expectations. Seeking help for such issues

is uncommon as the expression of negative emotions is frowned upon. Missing an outlet for the expression of emotions leads to a buildup of stress and casts a negative view on those with mental issues as well as the action of seeking professional help. While National Health Insurance (NHI) is readily available in Taiwan, individuals are not properly diagnosed and treated for Major Depression Disorder (MDD) (Chien, et al. 2007). The stigma of seeking professional help, misunderstanding of mental wellbeing treatment in general, and the cultural perception of self-expression has caused seeking professional help as a last resort (Lin, 2001).

**2.3 Chatbot History**

Chatter robot, more colloquially known as chatbot is a system that responds to users' interaction by simulating conversation either through text or voice. The first chatbot was developed in the 1960s with MIT's ELIZA computer program. Developed by Joseph Weizenbaum, ELIZA mimicked natural conversation by employing pattern matching and rule-based scripting. The most famous script that ELIZA ran was DOCTOR which simulated a psychotherapist session between it and the user. Recent chatbots employ Natural Language Processing (NLP) to understand specific contexts of the user's inputs.

More and more, we see the adoption of conversational agents in different mediums and varieties ranging from intelligent personal assistants such as Siri and Google Assistant to social robots like Pepper and Sophia to automated chatbots used in social media and customer support. In 2010, Siri was launched by Apple for iOS, which paved the way for the modern AI bots and IPA(Intelligent Personal Assistant). In 2016, Facebook further opened up the chatbot space by making chatbot development accessible for its social media platform. Line, the largest instant messaging platform in Taiwan (Statistica), officially enabled the development of chatbots in the same year. With the advancements we see in conversational interactions, conversational agents stand to provide more natural interaction methods for

users to interface with technology.

### 2.4 Chatbot in Mental Health

For chatbots in the field of mental health, a commercially available fully conversational chatbot using CBT methods named “Woebot” was developed in 2017. The research showed that the use of Woebot helped reduce anxiety for the participants (Fitzpatrick, 2017), as based on participant results of answering the General Anxiety Disorder 7 questionnaire (GAD-7), a commonly used questionnaire for measuring anxiety. TeenChat, a chatbot system to sense and release stress, was developed in China in 2015 which also showed promising results from computer-assisted mental treatment (Huang, 2015). Overall, chatbot use in mental health applications shows promise as conversational chatbots could provide unique assistance to mental health patients such as being available all the time, rapid accessibility, and avoidance of mental illness stigma.

Noteworthy is also the fact that as the current generation of adolescents has grown up with access to the internet and smartphones, they are more likely to be receptive to mobile health applications (mHealth) interventions, including chatbot interventions (Do et al., 2018; O'dea et al., 2018). Indeed, prior studies such as the aforementioned Woebot concluded that young adults using their intervention chatbot scored higher in satisfaction compared to a controlled group that relied on digital print material (Fitzpatrick, 2017). Another chatbot study (Crutzen et al. 2011) in regards to answering adolescents on topics related to sex, drugs, and alcohol also reported that adolescents preferred interfacing with chatbots in terms of the five features listed: anonymity, conciseness, ease of use, fun, quality and quantity of information, and speed compared to information lines, and search engines. From the research, it appears to hold true that the conversational characteristic of chatbots seems to be both favorable and effective in the domain of digital health.

The conversational interactions that chatbots use to approach mental health

interventions show promise in tackling mental health issues. As they mirror real-life interactions one would have with a therapist without the stigmatization of seeking help as well as being widely accessible to anyone with internet access, they stand to become the first step towards getting help (Kretzschmar et al, 2019).

However, there are still several issues that stand in the way of chatbots completely replacing traditional therapy. One theory as to why current solutions for mental health interventions have poor adherence is that there is a loss of quality of human interaction that traditional therapist-patient relationship still retains and mHealth apps lack (Fitzpatrick, 2017). While better personalization may help offer users more tailored treatment, this inability to re-create human interactions thus regulates current chatbots to more of a supplementary role rather than a complete substitute for mental health professionals.

Chatbots are still far from perfectly mimicking real life interactions and designers of chatbots draw upon several guidelines for conversational design such as *Conversational Design* (Hall, 2018) or *Designing Voice User Interfaces: Principles of Conversational Experiences* (Pearl, 2017) in order to create more human-like interactions. In Hall's *Conversational Design*, one maxim for human-human conversation that she draws upon for human-computer interaction is the maxim on being truthful. Hall translates this maxim in short for setting user expectations. As she puts it "Interactive truth demands a strong correlation between what the user expects and what the system offers" (Hall, 2018). This helps set the stage for the users to understand the extent of the chatbot's assistance and its capabilities.

# 3. Methodology

In order to fill the knowledge gap of what particular needs Taiwanese adolescents require for stress relief as well as what role chatbots can play in helping them, this study employs methods based on the ISO 9241 standard for human-centered design, which puts the needs and requirements of the user first when designing products and solutions. The human-centered design framework allows for an overall standard to be integrated into various design and development processes appropriate to the particular context of the design. To scope out high level requirements, project artifacts are created after which it is used to evaluate the design solution. In this section, the methods of this study will be described and the structure is as follows: an approach section describing what the approach of the study entails, a participants section detailing how participants are recruited and are involved within the overall study, a data collection section that talks about what methods of data collection is utilized for the study, and a data analysis section in which the analytical approach is explained.

## 3.1 Research Approach: Human-Centered Design Process

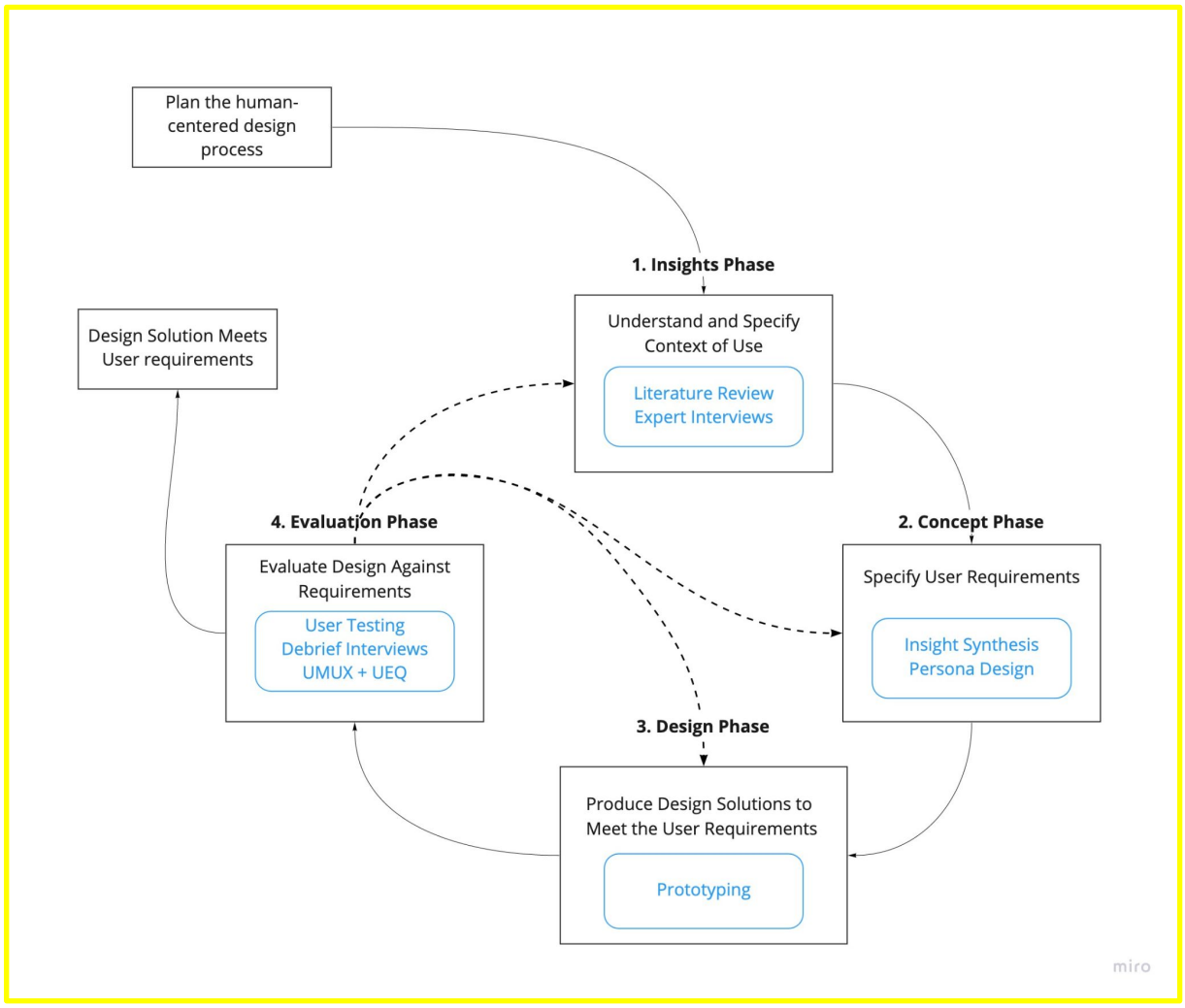
Based on the ISO standard 9241, this study follows a human-centered design process (ISO, 2010). The design process focuses on the human side of interactive design and seeks to put users' needs and experience at the forefront of the design process. Our users, as defined by the context of the study, are adolescents between the ages of 12 to 15, with a cultural background from Taiwan. Adjacent to the users is school counselors, as they are the ones to provide mental support for students in Taiwan, not school nurses.

The design process itself is split into four phases: a) an insight phase where research is done using qualitative research by the way of interviews with school counselors to understand the use case and context of Taiwanese adolescent needs and wants, b) a concept

phase in which the insights were synthesized into actionable concept that helps establish user requirements c) a design phase where a simple prototype was designed and d) an evaluation phase in which a focus group user tested the prototype and feedback was gathered in the form of questionnaires and interviews. An illustrative overview of the human-centered design process is provided in Figure 1. School counselors were involved in the early phases of the design process, providing insight for the insight phase and refining concepts created within the concept phase, whereas adolescents were involved in the evaluation phase, providing feedback and user testing.

**Figure 1.**

*Human-Centered Design process*




*Note.* Methods used are noted in blue within the overall Human-Centered Design process

### 3.2 Participant Recruitment

Five school counselors from three local middle schools in Taiwan were recruited through snowball sampling referrals early on to provide assistance in understanding what were the typical middle school students' needs and wants in regards to emotional support for stressful situations as well as how they, the counselors, approached counseling. Counsellors were given a 300 New Taiwan Dollar gift card (approximately €9) as an incentive for participating. All five counselors have had at least 2 to 10 years of actual counselling

experience. All five counsellors were involved in the insight phase and of those, four were involved in the concept phase, providing expert feedback and suggestions for the prototype.

From the four counsellors that assisted in the concept phase, three counsellors helped facilitate the recruitment of Taiwanese adolescent students from two schools respectively to participate in user testing. The counsellors invited from their own assigned pool of students a total of six adolescent students. All participation was voluntary and only the context of a chatbot that helps provide stress relief was provided to the adolescents beforehand. All adolescents participated in user testing to evaluate the prototype and were later on interviewed for feedback. Afterwards, without prior knowledge of an incentive, the adolescent students were given a gift card to the value of 200 New Taiwan Dollar (approximately €6) as thanks for participation.

### 3.3 Data Collection

Three methods were used for data collection in this study: interviews, expert feedback, and user testing. The language used in all data collection is Traditional Mandarin Chinese as all participants have native mastery of the language. All quotes from the participants are translated to English by the author for the purpose of this thesis.

#### *3.3.1 Interviews*

For the insight phase, individual interviews were conducted with five school counsellors, each approximately 1 hour in length, which were audio recorded with consent from the counsellors. The interviews were semi-structured in nature, with predefined key points to be discussed and leaving enough freedom for open-ended questions and answers. Key topics discussed in the interviews included what were the needs of Taiwanese adolescent students in terms of handling stress, how the counsellors help provide emotional support during counselling sessions, and the counsellor's opinions on adolescents seeking self help. The full interview plan can be found in Appendix A.

### *3.3.2 Expert Feedback*

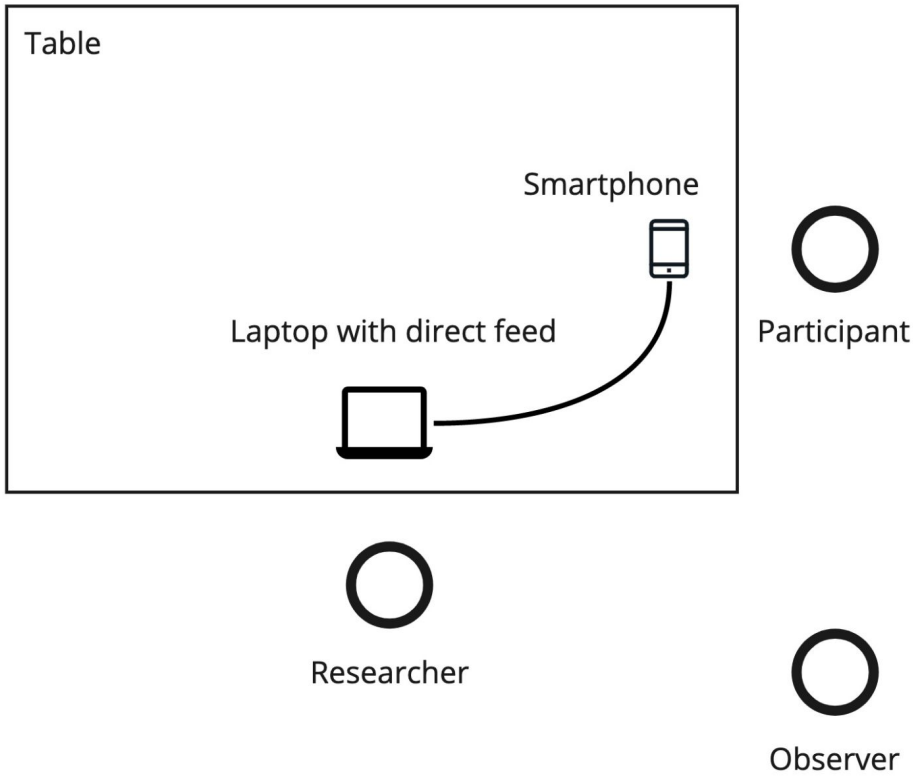
During the concept phase, a prototype of the chatbot was presented to 4 counsellors along with a dialogue map of the conversational interactions. Counselors were asked to provide feedback on what they thought of the chatbot in terms of the content provided, the dialogue language used, and the logic of the overall flow. Sessions were done individually, lasted 30 minutes to 1 hour each, and were unstructured. While no audio was recorded, remarks were transcribed and in particular, recommended changes to the wording and language of the dialogue were drafted by the counsellors.

### *3.3.3 User Testing*

In the evaluation phase, an iterated version of the prototype chatbot was used for user testing. 6 students participated in the user test which, though considered a low number, is deemed enough for the first user test in a series of iterations based on Nielson & Landauer's *A mathematical model of the finding of usability problems* (1993). The duration of individual user testing sessions were 45 to 60 minutes each. Testing was done in a Simple Single-Room Setup (Rubin & Chisnell, 2008), modified to include the counselor as a silent observer to ensure the student's safety and consisted of two chairs at a 90° angle, a separate chair off to the side, a desk and a smartphone device hooked up to a laptop via USB (see Figure 2).

**Figure 2.**

*User Testing Room Setup*




Before the user testing session began, a short brief of the user test was given to the participants verbally. The briefing describes the test as an exploratory research looking to discover what are the needs a chatbot should fulfill in order to help support middle school students in stress relief, allowing the participants to understand the context of the study. User expectation of the prototype is also reigned in beforehand during the briefing as more details on what a prototype entails were given to the participants.

After the briefing, the participants were presented with a user scenario of a 9th-grade middle school student who has not been to counselling before being stressed about an upcoming major exam, the Comprehensive Assessment Program for Junior High School Students (CAP). The use of this user scenario was a deliberate decision as it helps set up a common starting point for all participants to empathize with as well as protect the privacy of

the participants in question without invading upon their own issues. Participants were asked to perform 4 tasks during the scenario which were related to key features of the chatbot prototype (see Appendix B).

During the user testing, the participants were instructed to perform a *Relaxed Think Aloud* protocol (RTA) (Hertzum, 2016) while interacting with the prototype. RTAs differ from the more traditional *Think Aloud* of Boren and Ramey (2000) in that the researcher would probe the participant while the user testing was happening. Prompts such as "What do you expect to happen here?" and "How do you feel about this response from the chatbot?" helps to guide the participant to verbalize their immediate thoughts better. The decision to utilize RTA rather than the more traditional *Think Aloud* protocol was due to the belief that adolescent participants needed more guidance during the testing in order to express their thoughts on the prototype. Furthermore, is the context of dealing with the cultural background of Taiwanese adolescents, which based on Hofstede's cultural dimensions theory (2011) are generally less likely to speak out their opinions. Therefore, utilizing RTA for the testing of the chatbot for Taiwanese adolescent students would result in recommendations to redesign and overall user experience gathered more smoothly.

A short debriefing session followed after the completion of all 4 tasks in order to have the participants reflect upon their overall experience on using the prototype. A short debrief session which is a few open answer questions helps extract potentially missed insights that were not discovered during the user test as RTA is oftentimes overwhelming to participants that have not done a think aloud session before. The participants were asked in the debriefing session what they liked about the chatbot prototype, what they think could be improved, how they perceived the chatbot in terms of personality, and how effective they felt the chatbot was in explaining how school counselors can help.

After the debrief, the participants were asked to fill out two post-test questionnaires. The first is the Usability Metric User Experience (UMUX) (Finstad, 2010). The UMUX is a four item, 7-point Likert scale questionnaire designed to assess the application's perceived usability. Based on the more widely used System Usability Scale (SUS) (Brooke, 1986), UMUX is shorter yet still performs reliably to discover underlying perceived usability issues. In order to minimize misunderstandings stemming from cultural differences caused by direct translation of the UMUX, a cross culturally adapted version of the UMUX for Chinese users (Wang et al., 2021) was chosen for this study. Following the UMUX is a Chinese version of the User Experience Questionnaire (UEQ) (Laugwitz et al., 2008). The UEQ is a 26 question questionnaire that uses a seven-stage scale in order to reduce the well-known central tendency bias (Schrepp, 2015). Insights to 6 scales which are Attractiveness, Perspicuity, Efficiency, Dependability, Stimulation, and Novelty can be extracted using the UEQ to measure the overall user experience of the product. As part of the research question is in regards to the user experience of a stress relief chatbot, the UEQ questionnaire format was deemed reliable and relevant to the assessment of users' impressions and feeling after using the chatbot. Both the Chinese UMUX and the Chinese UEQ have been slightly modified from Simplified Mandarin Chinese to Traditional Mandarin Chinese as the official language used in Taiwan is Traditional Mandarin Chinese.

**3.4 Data Analysis**

Qualitative data gathered from the interviews with counselors and user testing debriefing were analyzed using a thematic analysis process as outlined by Braun and Clarke (2006). An inductive approach to the qualitative data was taken, in which the original data gathered in Traditional Mandarin Chinese was analyzed without modification. Analysis results are translated to English for the sake of this study. Looking at qualitative data gathered

in the insight phase, patterns were identified which were clustered and summarized into potential themes that would help drive the design process of the prototype.

Quantitative data will be derived from the UMUX and UEQ questionnaires. For the UMUX questionnaire, the median of the 4 5-point Likert scale questions will be calculated in order to understand the impression of the chatbot's usability. On the other hand, the UEQ questionnaire measures the participant's subjective 6 different scales related to attractiveness, pragmatic quality (Perspicuity, Efficiency, Dependability), and hedonic quality (Stimulation, Novelty) on a -3 to +3 scale with -3 being the most negative, 0 being neutral and +3 being the most positive response. A higher score would represent a high user experience.

# 4. Initial Results

The initial results chapter is structured below in three sections. The interview insights section discusses the key findings discovered through interviews with the school counselors. Next, in the system design section, the concepts and methods that lead to the development of the prototype are presented. Following this, the study presents the evaluation results of the user testing. All presented quotes from the interview are translated into English.

## 4.1 Taiwanese Adolescent Needs in Stress Support

In interviews with Taiwanese middle school counselors, the counselors highlighted that there is no single answer to what students need in terms of stress relief and support. Counselors mention that they experience a large variation in the type and severity of situations which makes their work more nuanced. Examples of stressor situations the adolescent students have brought up to the counselors include pressure of the Comprehensive Assessment Program for Junior High School Students (CAP) that happens during grade 9 which determines the high school a student could enroll in, interpersonal conflicts such as bullying and breakups as well as familial matters that range from severe issues like abuse and negligence to everyday conflicts.

> *"It's a very complicated question, what kinds of support I can give to students. And I can't really say as every student is an individual with their own problems. One student may be asked to come to me simply because their homeroom teacher thinks the student is being ignored, while another student is here because their parents are pressuring them to perform better in class. What I do as the counselor is let them know that they can tell me what they feel and I can listen and understand and we work together to solve the situation they are in" [Counselor 1]*

The counselors indicated that these aspects can be further complicated by the preconceived notion of counselling as most first time students have a negative perception of what it means to be counselled.

Counselors, therefore, describe that the first few sessions are dedicated to establishing trust and to building a relationship with the adolescent student for them to feel comfortable expressing themselves, as there could be an existing sense of hierarchy in the student's mind between them and the counselor. As indicated by Hofstede, the aspect of power distance exists in school settings. In high power distance cultures such as Taiwan, a teacher is to be respected or even feared at times. "Students in class speak up only when invited to; teachers are never publicly contradicted or criticized and are treated with deference even outside school." (Hofstede, 2010, p.84). Similarly, the teacher hierarchy is extended to school counselors as well.

> *"I try to build a sense that I'm not really their teacher, but rather someone who works at the school looking to help them. Because sometimes, they [the student] have a bit of resentment and fear that I'm working with the teacher who sent them to counseling in the first place. Oh and they are more closed-lipped about the teacher if the reason they were sent is listed on the note that they had a conflict with the teacher. So I have to talk to them and show them no, I'm not a teacher, I'm someone they can trust. And I do that first by simply talking to them about their interests and other easy topics and try to relate to them." [Counselor 3]*

In other words, first time students are intimated with the experience of being asked to participate in a counseling session, and counsellors seek to break down the power distance that is inherent between them and the student.

In addition to the counselors' efforts in establishing a friendly relationship with the student, the counselors also report the importance of educating what counseling entails. This is usually done through each school's promotion of counseling at the beginning of the year as well as mandated wellbeing courses that happen at least once each semester. However, the counselors also state that more can be done as despite the effort of schools to normalize seeking counseling, first time students still seem to lack information about counseling in general.

*"We do try to advocate for them to come to counseling if they are experiencing severe stress and other mental issues. We also have courses each month where we also teach them some mindfulness techniques such as breathing techniques and the like. And sometimes they work as we do get one or two students that come to our office asking for help. But I think we can do more and reach more of them."* [Counselor 4]

In this case, the needs of the adolescent student can be generalized as a) the need to be supported in a myriad of situations b) the need to be connected and empathized with and c) the need for information on counselling.

**4.2 Methodology of Counseling in Taiwanese Middle School**

Counselors describe that counselling mainly happens in Taiwanese middle schools via counsel requests by homeroom teachers. When a student's homeroom teacher believes that the student requires more professional mental support than the homeroom teacher can provide, they request intervention by the school counsellors and supply the counsellor with a short summary of the situation. The student is then asked to come to a counseling session where the counselor will try to discuss with the student their current situation. There, based on the issue in question, the counselor will try to assist the student in resolving their issue, and depending on the type and severity, overall assistance length may be only a few short sessions to 8 total sessions.

While it is largely dependent on the individual in question, the counselor's methodology of how to go about assisting the adolescent follows some common standards which also mirrors the needs of the adolescent students:

1) Building a safe space for expression and sharing.
2) Establish and maintain a friendly and confidential relationship not unlike that of a friend to decrease the power distance between the counsellor and adolescent.
3) Providing solutions either via actual and/or emotional recommendations.

The first standard helps address the wariness students have towards the counseling sessions and sets the room for the students to feel safe and trust the counselors. The second helps set the tone of all sessions so that the student continues to share and communicate with the counselor. The last standard addresses the stress itself either through resolving the stressor or through adjusting the student's emotional state.

**4.3 Seeking Self Help**

When discussing the counsellors' experience and opinions on current students self medicating through mental stress situations, the response given was that the counselors have not had much experience with middle school students in Taiwan trying to do so. Two major reasons were pointed out:

- reliance on social support
- low sensitivity towards their own situations and less alignment with their own feelings.

The first reason described by the counselors is that the students in Taiwan tend to rely more on informal support for help, whether it be friends or family. This is in line with prior research regarding attitudes towards help seeking within Chinese cultures. Boey (1991) reported that people in urban China were more likely to obtain help from social networks rather than professional helpers. Similarly, a survey conducted with 6498 Taiwanese regarding their medical seeking preference, mental health literacy, and attitudes toward mental illness came to the same conclusion that in Taiwan, informal help seeking was preferred at 64.2% compared to formal professional help seeking preference at 35.8% (Wu et al., 2014). Of the informal support preferred respondents, a majority (49.2%) preferred talking to their friends.

The counselors reported that when a stressful situation occurs in which the student is unable to regulate their stress by themselves, their first method of releasing stress is to vent it

with their peers, seeking emotional fulfillment and support through empathy. Many students seek empathy from communication with friends as they feel the need to be understood. It is less likely for them to proactively communicate with people above their “hierarchy” such as parents and teachers as they feel that there is a distance between them preventing them from truly understanding their feelings. This preference for social support has led to Taiwanese adolescents relying on social support rather than looking for ways to help themselves in stressful situations.

The second reasoning has to do with Taiwanese adolescents’ own sense of self. Counsellors interviewed have mentioned that compared to western counterparts, self expression is not as common as adolescents in Taiwan. Especially during the time of adolescent growth, Taiwanese students worry mostly about fitting in the norm. Coupled with the still maturing sense of self, self expression and self realization are not something Taiwanese adolescents strive to achieve, leading to a low sense of their own feelings and situation.

# 5. Implementation

In this section, the design of the prototype is described in 3 parts: a) technical design b) content design c) persona design. Technical design details the technical architecture of the chatbot, what platform was used to develop the prototype, and why. Content design describes three functions that address the discovered insights from interviews with counsellors, how they formed the content of the chatbot and how they are implemented. Persona design focuses on what personality and traits the prototype chatbot embodies.

## 5.1 Technical Design

The purpose of the preliminary prototype was to explore in what ways the needs of Taiwanese adolescents identified during the insight phase can be addressed and not meant to test the effectiveness of the chatbot on its users' stress levels. This meant that functionalities related to the needs should be somewhat complete, or "horizontally represented" as Rubin & Chisnell (2008) would call it. However as users needed to interact with each function to an extent in order for the study to understand how well the chatbot supports fulfilling the users' needs, a certain degree of "verticality" was also required. Google's Dialogflow platform (https://dialogflow.cloud.google.com/) was used to develop the prototype chatbot as the author has prior experience with the platform and also due to the fact that Dialogflow allowed for integration with Line, the most widely used social messaging application in Taiwan. Users navigate and trigger specific functions called intents in Dialogflow through the use of either predefined answers or freeform input. Dialogflow allows for the detection of intent based on trained phrases and keywords as well as looking at prior interactions' output context and responds with the correct text message within the dialogue tree.

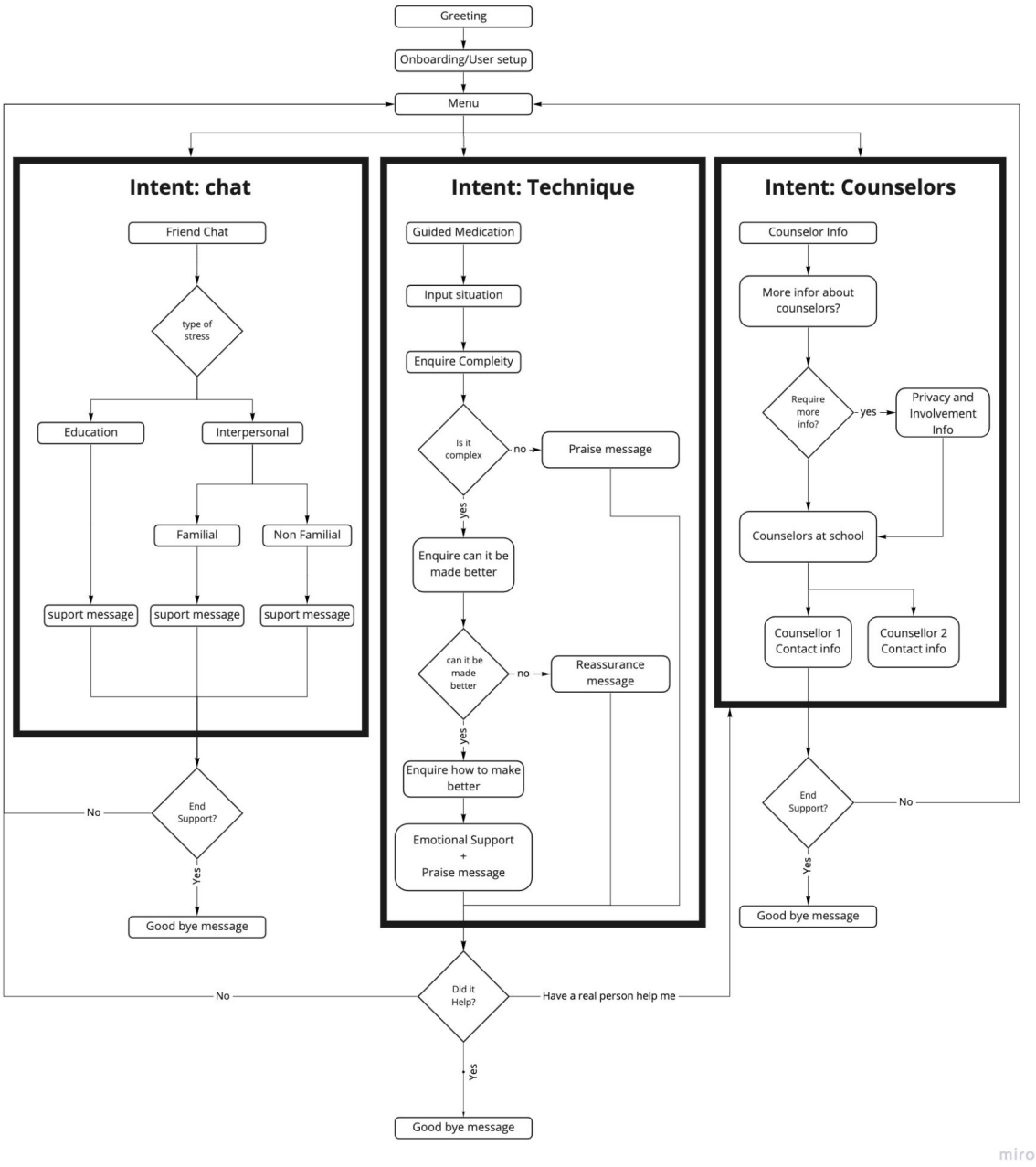
## 5.2 Content Design

Based on the findings of the interviews with counsellors, the content of the prototype strived to address the 3 needs described (varied situational support, connectedness support,

and counselling info support) in section *4.1 Taiwanese Adolescent Needs in Stress Support*. The standards used by the counselors were also taken into consideration in the content flow of the prototype. Finally in regards to self medication, as it was mentioned to be less likely to happen due to the cultural nature of the adolescent, a guided meditation function was also conceptualized. Content flow was mapped on Miro and Figure 3 shows the high-level content architecture tree of the content.

**Figure 3.**

*High Level Content Architecture Tree.*




As the prototype is aimed to be used as an exploratory artifact rather than a completed product, the extent of the chatbot content design was kept to only the most necessary focus. Therefore in the case of this study, only three dialogue trees that were based on the discovered user needs and counselor methodologies were fleshed out.

### *5.2.1 Intent: Chat.*

The user insights from the counselors mentioned that building a relationship between them and the students was vital in opening up the space for effective counseling. Not only was it necessary for the counselors to establish that they were trustworthy for adolescent students to share their feelings with as a friendly advisor but also crucial as a wrong impression could create a power distance between them and the students. This is also further supported by the insights from the topic of self help as the adolescents' reliance on their own social network means that to them, a necessary part of stress relief and support lies in having interactions with their friends and being supported through empathetic dialogue and acknowledgment to the adolescent’s stress and plight.

In order to address the needs of providing connection and empathy, the chatbot prototype mirrors the relationship and conversational pattern a Taiwanese adolescent student would have with a close friend. This emotional support is provided all throughout the chatbot's conversation style (further discussed in 5.3 Persona Design) but is especially evident in the Intent: Chat functionality. The function Chat allows the user to interact with the chatbot in a short conversation where the chatbot asks what stressful situation the adolescent is experiencing and based on the response, would detect what type of stressor is causing the situation and provide a relevant supportive message (see Figure 4). The support messages provided in response were based on categories of support messages theorized by Kindness et al. in *Designing emotional support messages tailored to stressors* (2016) such as Emotional Reflection, Directed Action, and Reassurance, which were later on validated and revised tonally by the counselors to be similar to their counseling techniques during the expert feedback round.

**Figure 4.**

*Screenshots of the Chat Function*


*Note.* In the example, the chatbot provides a relevant support message based on the detected stressor. Original Traditional Mandarin Chinese version (left) was used for testing, and the translated version (right) was made for this study.

### *5.2.2 Intent: Technique.*

To see whether Taiwanese adolescent's indifference towards self help could be assuaged, a different approach was taken to help users deal with stress and negative emotions within the function intent techniques. Like other established chatbots for mental health, the role of the chatbot becomes more of a counselor or professional, assisting in providing

techniques that users can use to help care for their mental health. The prototype chatbot design adapts techniques that the counselors in Taiwan implement during their sessions to help the adolescent students break down situations into smaller parts and provide them with a more balanced view of the overall situation. Hence, established methods used by the counselors were adapted to suit a conversational user interface in terms of structure and phrasing.

The prototype chatbot initiates by asking the user to talk about what their current stressful situation is and guides them to reflect upon their thoughts of the matter, whether they felt it was complex, whether the situation could get better and whether the user currently has any thoughts on how they could achieve that ideal solution. Depending on the response of the user, the chatbot will respond with an appropriate emotional valence and message for emotional intervention using similar support message techniques found in the chat function (see Figure 5). Similarly, during the expert feedback round, the counselors helped review and revise the messages to be tonally closer to what they use in their counselling sessions.

**Figure 5.**

*Screenshots of the Technique Function*


*Note*. In the example, the chatbot guides users to discuss and reflect on a stressful situation and respond with messages that employ emotional support message categories of Emotional

Reflection, Reassurance, and Emotional Advice. Multiple screens are stitched together. Original Traditional Mandarin Chinese version (left) was used for testing, and the translated version (right) was made for this study.

### *5.2.3 Intent: Counsellor.*

The last function, intent counsellors seeks to address the need of information to counselling discovered in the interviews. Considering that situations that are more serious or cannot be solved with what the chatbot can provide assistance for may happen, support from real professionals will be required. In these situations, the chatbot should incentivize and provide the users with access channels to proper mental health support from professionals. However as discussed prior, there is still a huge barrier in the form of stigmatization to seeking professional help for mental issues in Taiwan. As Taiwanese adolescents' perception of counselling was that it was mainly for the severely troubled person, there exists a need to normalize and open up channels for seeking help. To tackle and reduce the barrier of counselling, providing information of what counselling is and reassurance of the confidential nature of counselling to the uniformed via a friendly manner is especially important.

To help support this need, users are prompted to seek professional assistance in various points of the interaction, such as during the onboarding when explaining how much assistance the chatbot can provide as well as at the end of the technique function when the chatbot enquires whether it has helped in providing enough stress relief for the user. Within the overall structure of the chatbot, the support for information about counselling is softly suggested, by first asking the user if they want a real person to intervene and help and then opening up the dialogue about counsellors at school by providing information on what counseling entails, the confidentiality of counselors' work and details on getting in contact with them such as office location, email and phone number to use (see figure 6).

**Figure 6.**

*Screenshots of the Counselors Function*


*Note.* In the example, the chatbot provides users information on what counselors do and their confidentiality nature. Original Traditional Mandarin Chinese version (left) was used for testing, and the translated version (right) was made for this study.

### 5.3 Persona Design

According to Fitzpatrick et al. (2017), the relationship between a person and a chatbot is a crucial aspect in how effective a chatbot can help in regards to mental illnesses. Likewise, the insight phase also underpins that building a relationship between them and the Taiwanese adolescent students is essential and the first step they took in order to ensure that the student is willing to trust and share their thoughts and worries with them.

The insights, however, stress that despite the efforts of the counselors, there is still a slight barrier to establishing a completely natural relationship between the counselors and Taiwanese students. Power distance still exists between the counselors and their students, which makes it complex to broach into sensitive situations. To deal with this, care was taken in the design of the chatbot's personality so that there was no preconceived "hierarchy" for the adolescent users. The chatbot's persona, therefore, was modeled to be a friend with slightly more experience with the subject matter, dealing with stress, similar to an upper classmate.

Linguistic idiosyncrasies of the Taiwanese adolescents were mirrored and used by the chatbot to accentuate the perception of the chatbot as a friendly peer. The dialogue, language used and expressions were all explicitly chosen to help communicate and to reflect a sense of empathy and companionship to the user. The personality design reflects the need of the chatbot in supporting the user during stressful situations, which is why the functions and dialogues of the chatbot focused less on the actual mental illness and distress, but rather on providing supportive messages to the users and to encourage Taiwanese adolescents to seek professional help in the form of counselors all through the frame of a friendly peer relationship between the chatbot and the user.

# 6. Evaluation Results

In this section, the findings of the research will be discussed. First the qualitative findings from both the Relaxed Think Aloud protocol as well as the debrief interviews will be presented, followed by the qualitative results of the data from the UMUX and UEQ questionnaires. An analysis of the overall results will be discussed and addressed in the Discussion chapter. The summary of both questionnaire results from Google Forms can be found in Appendix C.

## 6.1 Qualitative Results

Six Taiwanese adolescent students that participated in the user tests helped provide feedback on their user experience interacting with the prototype chatbot. The qualitative findings based on their comments (noted as P1 to 6) during the Relaxed Think Aloud and the debrief interviews are as follows:

**Chatbot as a friend helps support Taiwanese adolescents in self expression of problems:** All six participants remarked that the chatbot persona was quite evident and helped them be more willing to talk and vent about emotional issues. When asked why, several reasons were elaborated on such as feeling no pressure to respond as it was not a real person, the perception of talking to a friend that has a bit more experience in life, and a feeling of companionship. The participants described the chatbot as someone that they could open up to.

> "*When I see how the chatbot responded to me, I get the feeling that I could tell it things that worry me about without having to care about being judged. Like, I probably would share more with a friend the same age as me than with a counselor." [P3]*

This view of the chatbot as a friend as well as the comment suggests that even in normal counseling sessions, the adolescent can feel judged when talking to the counselor. The role of a relatable friend, therefore, seems to be quite important in fostering open dialogue.

**Supportive messaging should be more layered and humanlike:** During the function of Intent: Chat, the chatbot detects the stressor and sends a message based on the type of stress immediately. While the support message itself was liked by five out of the six participants, all six participants felt that the immediateness of this made the support message feel "superficial".

> *"It [the support message] is good I think, like something my friend would maybe say, but I wish it was more deep, like it can ask me more, instead of immediately giving me this support message." [P4]*

This lack of 'deepness' as the participants mentioned, seems to be largely attributed to both the speed of response as well as the lack of follow up dialogue. The lack of follow up dialogue also meant that participants felt that the chatbot was not as interested in interacting or that the chatbot was in a hurry to give support.

> *"While I liked that the response is fast, it's a bit to the point. Makes it seem like I'm talking to a friend that is just trying to quickly end the conversation. Kind of like me when I don't really want to know the details." [P6]*

Several participants commented that the response feels a bit unnatural because there was no delay between messages. However, it should be mentioned, Dialogflow, does not support as of yet, artificial delays, otherwise it would have been implemented in the prototype.

**Expectation of credibility and professional background in guided self medication of stress:** Four out of the six participants indicated that they felt they would not use the

guided stress relief method on their own if given the choice. When asked to elaborate on this, the participants mention that they don't believe it would help them as their perception of the chatbot was that it was not a counselor or a professional.

> *"The problem I have with this [guided techniques] is that I feel like it would be very complex. It is just asking me to reflect, but I want it to provide actual solutions to my problems and I think only the counselor can do that."* [P4]

From the comments, it could be inferred that the prototype itself was not designed adequately enough to give proper advice on how to deal with stress and that by involving professionals in the design of this could make the participants feel differently to being provided a stress relief technique. Consequently, the chatbot persona of a friend worked against its credibility:

> *"I dunno, I feel like if a friend told me to do this, I wouldn't really believe it to work. I mean, they are supposed to be the same age as me, right? So how would they know?"* [P2]

**Awkwardness in approaching counselors:** As one of the aims of the prototype was to see if the negative bias about counselling could be addressed through the use of a chatbot, the participants were asked to approach this as though they have not had counselling before in order to understand their opinions on being provided info on privacy and how to get in touch with counselors. The participants reacted to the information provided by the chatbot in the role of a friend on counselling and privacy fairly well, noting that it made them more aware of the confidentiality of the counselling sessions, and would thus be more trusting in seeking help from counselors.

> *"I think that the information provided is quite good and if I haven't been to counselling before and a friend recommended it to me, it would help me trust it [the school counselor office] more."* [P5]

However, in regards to the contact information provided by the chatbot, five out of the six participants felt negative about it as they felt awkward reaching out to a stranger. They mentioned that they would not seek to use the information the chatbot provided to get in touch with counselors as they don't know them enough.

> *"I actually don't think anyone would use this to contact counsellors. Maybe if I was already familiar with the person listed, but otherwise, I would not contact them....I feel awkward enough having to talk to unfamiliar people."*[P2]

### 6.2 Quantitative Results

**UMUX Results**. The first part of the debriefing questionnaire is the Chinese UMUX questionnaire, used to determine the perceived usability of the prototype chatbot. Each item is scaled on a 7-point scale with 1 being "Strongly Disagree" and 7 as "Strongly Agree". The scores are then recoded as per Finstad's (2010) formula: "Odd items are scored as [score − 1], and even items are scored as [7 – score]" in order to remove positive/negative keying of the items, similar to SUS. As the number of participants is lower than 30, the use of population standard deviation was chosen. The descriptive statistics of the UMUX results can be found in Table 1.

**Table 1.**
UMUX Response Results (n=6)

| Question | Tone | Mean | Standard Deviation |
|---|---|---|---|
| This product's capabilities meet my requirements. | Positive | 4.666 | 1.211 |

| | | | |
|---|---|---|---|
| Using this product is a frustrating experience. | Negative | 3.333 | 2.658 |
| This product is easy to use. | Positive | 4.833 | 1.471 |
| I have to spend too much time correcting things with this product | Negative | 2.833 | 1.471 |

From the results, it can be inferred that the prototype performed above average in the perceived usefulness and usability department as seen in questions 1 and 3 respectively. However, as seen in questions 2 and 4, the participants felt rather negatively about the prototype in terms of operating the prototype, possibly suggesting that in terms of technical perception, the prototype performed not up to expectations.

To benchmark the prototype, an overall score was calculated using the formula provided by Finstad (2010) to generate a usability score comparable to the SUS's output. The results are shown in Table 2.

**Table 2.**
Overall UMUX Response Results in SUS scores (n=6)

| Mean | Standard Deviation |
|---|---|
| 65.277 | 21.191 |

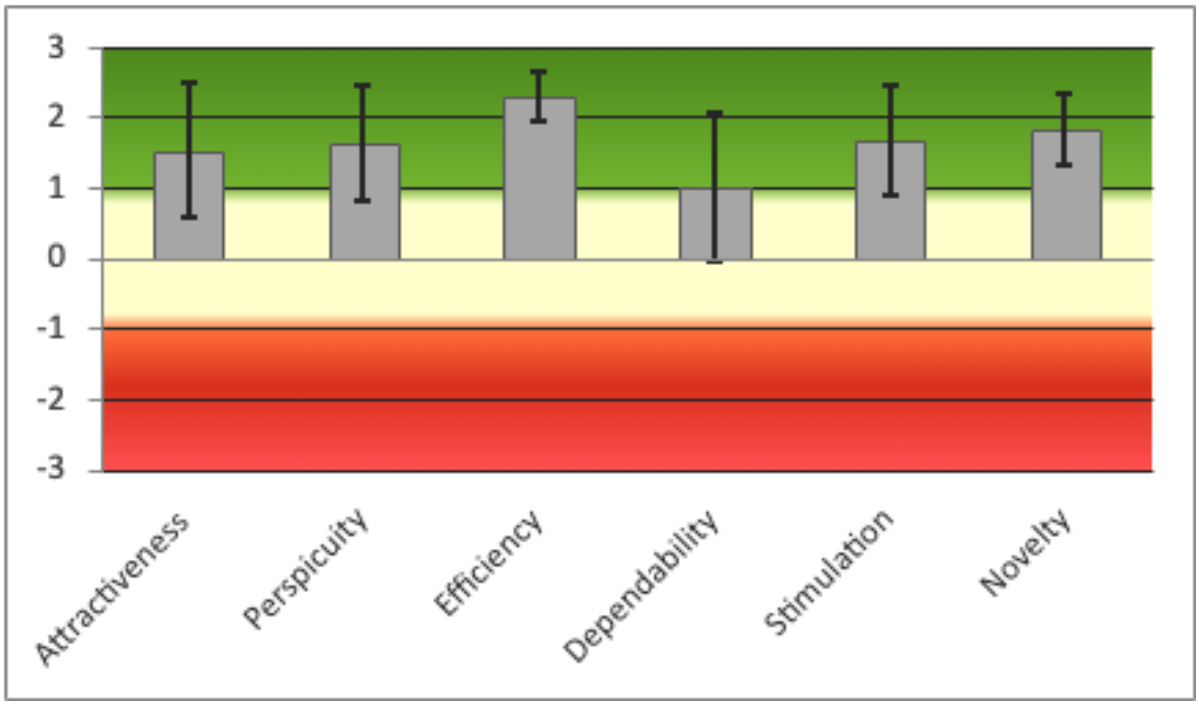
Overall the mean score suggests that usability-wise, the chatbot scored a 65.277 in SUS format. This is comparable to a C- grade, as based on Sauro and Lewis' (2016), proposed curve grading scale (CGS) for SUS scores where the SUS scores range from F (absolutely unsatisfactory) to A+ (absolutely satisfactory). Sourced from 446 studies and over 5,000 individual SUS responses, Sauro and Lewis found the average mean score of the SUS

to be 68 with a standard deviation of 12.5. Moreover, within the study, Sauro and Lewis delineated the scores based on the categories of product/service surveyed into 8 different types: B2B, B2C, Web, Cell, HW (Hardware), Internal SW (Software), IVR (Interactive Voice Response), and Web/IVR, each with their own mean score and standard deviation. From this benchmark, the chatbot prototype performed slightly below the average of the global product/service surveyed and slightly better than the benchmark for Web/IVR (mean score of 59.2). In line with the individual item scores, an explanation for the lower score could be that as the chatbot's technical functionality is not vertical enough, the participants felt that usability wise, it performed lower than expected.

**UEQ Results.** The second questionnaire used in testing was the Chinese UEQ. An overview of the processed data can be found in Appendix D. The UEQ questionnaire had the participants answer 26 items that were used to measure 6 different dimensions of the User Experience Questionnaire (UEQ) which are: *Attractiveness, Perspicuity, Efficiency, Dependability, Stimulation,* and *Novelty*. Figure 7 shows the overall results for the six UEQ dimensions mentioned.

**Figure 7.**

*Overall results of the User Experience Questionnaire (UEQ) in accordance with the six dimensions.*




All 6 dimensions were rated in the positive (0 to +3), with *Efficiency* performing the highest with a mean value of 2.292 and variance of 0.21, while *Dependability* performed the lowest with a mean of 1.00 and variance of 1.73. The results suggest that the prototype was perceived as efficient but undependable which is in line with what the UMUX results suggested. Of the 26 different items, the 3 top and 3 bottom performing aspects of the prototype chatbot are listed in table 3. The top performing item fast/slow received a mean score of 2.8 with a very low variance of 0.2. The second top performing item, organized/cluttered, scored a similarly high mean of 2.7 with a variance of 0.7 and following that, the item valuable/inferior scored a mean of 2.5 with a low variance of 0.3. This would suggest that the prototype chatbot is perceived as *fast*, *organized,* and *valuable*. This is contrasted with the bottom performing items that suggest that the prototype chatbot is somewhat *not understandable*, *obstructive,* and *unpredictable*. However, despite the lower mean scores of the items, all 3 still performed above or equal to a neutral mean score (0.8). That along with the high variance of the response for these items may suggest that more data may be required to get a better grasp of the perceived user experience of the prototype.

**Table 3.**
Top 3 and bottom 3 performing UEQ Responses

| Item | Mean | Variance | Standard Deviation | Scale |
|---|---|---|---|---|
| fast / slow | 2.8 | 0.2 | 0.4 | Efficiency |
| organized / cluttered | 2.7 | 0.7 | 0.8 | Efficiency |
| valuable / inferior | 2.5 | 0.3 | 0.5 | Stimulation |
| not understandable / understandable | 1.0 | 5.2 | 2.3 | Perspicuity |
| obstructive / supportive | 0.8 | 5.8 | 2.4 | Dependability |
| unpredictable / predictable | 0.8 | 3.8 | 1.9 | Dependability |

# 7. Discussion

The two research questions this thesis set out to answer were: What roles should a chatbot for stress relief support for adolescents in Taiwan fulfill to address their specific needs and What makes a good user experience for Taiwanese adolescents using a text based chatbot for stress relief? The insights and implications discovered are discussed below in two sections. The first half of the section details the user needs chatbots in the mental health space within Taiwanese culture should fulfill as well as what roles the chatbot should perform in. The second part of the discussion section describes user experience insights that were uncovered by the research and the implications for chatbot design in similar domains.

## 7.1 User Need Implications

### *7.1.1 The Role of "Friend" for Chatbots.*

One of the major concerns and focus during the research for designing a chatbot for Taiwanese specific audience was regarding whether relationship would play a major role in the perception of usefulness. As touched upon in *2.1 Stress and Culture,* an individual's cultural background can affect their expression of emotion. Within what Hofstede (1989) would term as a collective culture, in order to preserve the harmony of the society they live in, people form hierarchies both physically and mentally. As the subject of study was Taiwanese adolescents, the chatbot was designed to be perceived as a friend within a similar hierarchy rather than a counselor who was perceived as someone at a higher hierarchy, to reduce power distance. The interviews during the research and as well as the results of the evaluation phases revealed having the chatbot be viewed as a friend was beneficial to getting the users to share and express their emotions. The main aspect emphasized by the participants was that in this relationship, trust towards the chatbot is more easily built, suggesting that the role of a "friend" could be beneficial towards getting Taiwanese adolescents to open up to chatbots as can be seen from the interview quoted in *6.1 Qualitative Results*.

### *7.1.2 Perception of Professionalism.*

On the flip side, in order to motivate Taiwanese adolescents to try self medication for stress relief, the chatbot needed to be perceived as someone with a higher hierarchy and with credibility. As noted in *4.1 Taiwanese Adolescent Needs in Stress Support,* power distance that exists between a teacher and students extends to school counselors as well. While in face-to-face counseling sessions, counselors work to diminish the distance, it would seem that having at least some form of power distance is still beneficial to Taiwanese adolescents in regards to chatbot perception. The user test results indicated that the more friend-like the chatbot was, the less it was perceived as credible and professional, which decreased the willingness of Taiwanese adolescents to try stress relief methods. Participants reported that they felt weird that a friend was trying to provide counselling techniques. They pointed to the language used, as well as emojis and stickers to work against the credibility of the chatbot. This insight could point to a need to develop a more professional chatbot persona in order to affect Taiwanese users' perception of usefulness for the chatbot.

### *7.1.3 Addressing Bias through Familiarity*

As stigma towards mental illness and seeking professional help is especially prevalent in Taiwan, one major function in the chatbot was to equip users with information on counseling to help demystify and promote help seeking as well as provide a channel for Taiwanese adolescents to easily get in touch with counselors. The evaluation of the user testing results suggests providing information on what counseling entails as well as the confidentiality of sessions is looked upon favorably and would likely alleviate some of the negative bias towards help seeking. However, simply providing a channel to contact counselors is not viewed as favorably as the participants remarked that sharing their worries is awkward if they did not know the counselors beforehand. This could suggest providing

more information on specific counselors that help foster familiarity could improve Taiwanese adolescents' willingness to seek help.

## 7.2 User Experience Implications

### *7.2.1 Human-like Interaction.*

From the evaluations of user testing, the most remarked experience in using the chatbot was the human-like interaction done via texting. In all user testing sessions, participants have remarked that they liked interacting with the chatbot as if it was a friend, making it easy for them to chat with the chatbot. In addition, the use of emojis, stickers, and linguistic idiosyncrasies helped establish an engaging and helpful relationship between the chatbot and the user. The downside of this is that participants started to expect more natural human-human interactions. as they tried to write and converse freely with the chatbot. This caused the chatbot to not understand the participants' input and led to general disappointment of the chatbot's capabilities which is reflected in the results of the fourth item on the UMUX questionnaire as well as the dependability scale of the User Experience Questionnaire. With the advance in artificial intelligence, machine learning, and natural language processing, future user expectations towards human-like interaction could be better fulfilled.

### *7.2.2 Non Human-like Function.*

An interesting insight learned from this research is that not all humanistic characteristics of conversations are appreciated when interacting with a chatbot. As indicated in the evaluation results, the participants appreciated that the chatbot responded quickly as it felt attentive to their communication. This is also backed up by the results from the User Experience Questionnaire with *fast* being the top performing item. Furthermore, while the chatbot is perceived as a friend, the understanding that it is not a real person also helps in allowing users to share what they think and feel without feeling judged, which is in contrast to human-human communications where they have to adjust what they say to fit within the

social norms. This could suggest that the technical idiosyncrasy of chatbots not being perceived as a human agent might end up as a positive.

# 8. Conclusion

As the popularity of chatbots grows in all domains around the world, the focus on the particularities of cultural backgrounds and how it affects chatbot design should not be ignored and be investigated more thoroughly. This study, therefore, aims to contribute and extend the research on how to go about designing a chatbot but in the context of the adolescent population in Taiwan and how to approach helping them in stress relief. An exploratory study was done to identify a list of key needs that are essential to Taiwanese adolescents in achieving some form of stress relief, as well as learn what factors shape a good user experience in this context. The results of the thesis will be concluded in this section followed by the limitations of the research and what future research could be done to extend knowledge in the design of chatbots.

## 8.1 Research Summary

Using a Human-Centered Design approach, Taiwanese middle school counselors were interviewed for their insights in counselling Taiwanese adolescents in school as well as their perspectives on how to approach designing a stress relief chatbot. A conversational chatbot was then conceptualized and developed with Google Dialogflow which was evaluated using a Relaxed Think Aloud protocol for qualitative data and two quantitative Likert-scale questionnaires. The results helped provide insights for the two research questions posed in the introduction section:

***What roles should a chatbot for stress relief support for adolescents in Taiwan fulfill to address their specific needs?***

The relationship a chatbot builds with Taiwanese adolescent users determines what needs can be fulfilled. On one hand, a chatbot with a relatable and friendly personality is perceived as less professional but more conducive to emotional expression. On the other hand, a more professional but less relatable is viewed as more credible and invites users to

trust in stress relief techniques. This is due to how culture influences Taiwanese adolescent perceptions of competency. Those higher in the social hierarchy are perceived as more competent. In this case, depending on what the chatbot seeks to do, different roles need to be presented to the user. Furthermore, providing information on counselling through the role of a friend was remarked to be effective as the information provided is perceived as reliable and trustworthy.

***What makes a good user experience for Taiwanese adolescents using a text based chatbot for stress relief?***

The results of this research pointed to two contradictory expectations that provide a good user experience for Taiwanese adolescents. Users expect chatbots to perform human-like complex multilayered conversations that foster deeper communications and empathy. At the same time, the speed in receiving a response and the non-human nature of the chatbot is also seen as valuable and conducive to a good user experience. Participants point to these qualities as providing a sense of attentiveness and non-judgmentalness. This points to the need for efforts of improving conversational interaction to be more humanistic without sacrificing technical aspects of chatbots that differentiate it from human to human conversations.

**8.2 Limitations**

Due to the exploratory nature of the research, this thesis faces a couple of limitations. First, a limitation to this study is in regards to the observer bias in qualitative research. As qualitative research was performed by a single researcher, it is possible that observer bias exists to some degree. As there was not a second observer beside the author to help alleviate observer bias, quantitative research from questionnaires was also employed. The hope is that with the inclusion of quantitative results observer bias in interpreting data can be somewhat counteracted.

Moreover, while the research aims to provide new insights towards how to design chatbots for Taiwanese adolescent students, the study features a very limited sample size of 6 participants overall. While industry insights have claimed that a sample size of 5 is adequate enough for usability testing (Nielson & Landauer, 1993), the study done here involves more than just usability and thus statistical proof of the results cannot be fully claimed, rather the aim is in identifying general tendencies and perceptions. Moreover, the adolescent participants were in general well adjusted, having gone through counseling already. If a more troubled participant group were instead to be involved, the results could differ from the ones presented in this study. Therefore care should be taken when using the results to generalize for the overall adolescent group in Taiwan. In general, further research should be conducted to refine, validate, and expand the results found in this study.

### 8.3 Future Work

Due to time and technical limitations, not all aspects of chatbots were researched in this study. In this section, we discuss some of the ways that the research can be extended.

In general, more effort and time could be spent in the design of the overall conversations, extending its ability to interact with the user more naturally. Methods that could be done include implementations of small talks as well as more complex dialogues that overall improve the humanlike interactions of the chatbot to see if the results show further proof of Taiwanese adolescents' preference towards friend-like chatbots indeed correlates to willingness to share personal issues. In this study, a friend-like persona was chosen to be the overall impression Taiwanese adolescents could most likely trust. While the results show promise towards this assumption, split A/B testing using two different personalities could also shed more light upon the same subject.

Additionally, as touched upon in the limitations section, a way to extend this research is through additional research with a more troubled group of participants. The inclusion of

more participants of varying degrees of stress could allow for an even fairer evaluation of the relative perceived benefit of the chatbot's role. Empirical research such as randomized controlled trials or longitudinal studies could be done to test the efficacy of the chatbot's support for stress.

In terms of research related to the effects of culture in chatbots, this study draws on only one specific user group and culture: Taiwanese adolescents. The research into cultural impacts on chatbots has wide reaching implications towards chatbot development and more research is recommended for the subject. While this study imparts valuable insights for this particular context and culture, future research should be conducted for other cultures and age groups in order to extend the body of knowledge about the role culture plays on human–chatbot interaction.

## Reference

American Psychological Association. (2013). Stress in America™ 2013 Highlights: Are Teens Adopting Adults' Stress Habits? *Stress in America 2013 Highlights: Are Teens Adopting Adults' Stress Habits?* https://www.apa.org/news/press/releases/stress/2013/highlights.

American Psychological Association. (2018). Stress in America: Generation Z. Stress in America™ Survey. *Stress in America™ Generation Z.* https://www.apa.org/news/press/releases/stress/2018/stress-gen-z.pdf.

American Psychological Association. (2020, October). Stress in America™ 2020. *Stress in America™ 2020: A National Mental Health Crisis*. https://www.apa.org/news/press/releases/stress/2020/report-october.

Boey, K. W. (1999). Help-Seeking Preference of College Students in Urban China After the Implementation of the "Open-Door" Policy. International Journal of Social Psychiatry, 45(2), 104–116. https://doi.org/10.1177/002076409904500203

Boren, T., &; Ramey, J. (2000). Thinking aloud: reconciling theory and practice. *IEEE Transactions on Professional Communication*, 43(3), 261–278. https://doi.org/10.1109/47.867942

Brooke, J. (1996). SUS: A 'Quick and Dirty' Usability Scale. *Usability Evaluation In Industry*, 207–212. https://doi.org/10.1201/9781498710411-35

Chien, I.-C., Kuo, C.-C., Bih, S.-H., Chou, Y.-J., Lin, C.-H., Lee, C.-H., & Chou, P. (2007). The Prevalence and Incidence of Treated Major Depressive Disorder among National Health Insurance Enrollees in Taiwan, 1996 to 2003. *The Canadian Journal of Psychiatry*, *52*(1), 28–36. https://doi.org/10.1177/070674370705200106

Crutzen, R., Peters, G., Portugal, S., Fisser, E., & Grolleman, J. (2011). An Artificially Intelligent Chat Agent That Answers Adolescents' Questions Related to Sex, Drugs, and Alcohol: An Exploratory Study. *Journal of Adolescent Health*, 48(5), 514–519.

https://doi.org/10.1016/j.jadohealth.2010.09.002

Crystal, D. S., Chen, C., Fuligni, A. J., Stevenson, H. W., Hsu, C.-C., Ko, H.-J., … Kimura, S. (1994). Psychological Maladjustment and Academic Achievement: A Cross-Cultural Study of Japanese, Chinese, and American High School Students. *Child Development*, *65*(3), 738. https://doi.org/10.2307/1131415

Do, T. T. T., Le, M. D., Nguyen, T. V., Tran, B. X., Le, H. T., Nguyen, H. D., … Zhang, M. W. B. (2018). Receptiveness and preferences of health-related smartphone applications among Vietnamese youth and young adults. *BMC Public Health*, *18*(1). https://doi.org/10.1186/s12889-018-5641-0

Finstad, K. (2010). The Usability Metric for User Experience. *Interacting with Computers*, 22(5), 323–327. https://doi.org/10.1016/j.intcom.2010.04.004

Fitzpatrick, K. K., Darcy, A., & Vierhile, M. (2017). Delivering Cognitive Behavior Therapy to Young Adults With Symptoms of Depression and Anxiety Using a Fully Automated Conversational Agent (Woebot): A Randomized Controlled Trial. *JMIR Mental Health*, *4*(2). https://doi.org/10.2196/mental.7785

Hall, E. (2018). Conversational Design. O'Reilly (1st ed.). A Book Apart. https://learning.oreilly.com/library/view/conversational-design/9781492017950/

Hertzum, M. (2016). A usability test is not an interview. *Interactions*, 23(2), 82–84. https://doi.org/10.1145/2875462

Hofstede, G. H., Hofstede, G. J., & Minkov, M. (2010). Cultures and organizations: software of the mind. ProQuest eBook Central (3rd ed.). McGraw-Hill. https://ebookcentral.proquest.com/lib/tampere/reader.action?docID=6262353

Hofstede, G. (2011). Dimensionalizing Cultures: The Hofstede Model in Context. Online Readings in Psychology and Culture, 2(1). https://doi.org/10.9707/2307-0919.1014

Hsieh, H.-F., Zimmerman, M. A., Xue, Y., Bauermeister, J. A., Caldwell, C. H., Wang, Z., & Hou, Y. (2014). Stress, active coping, and problem behaviors among Chinese adolescents. *American Journal of Orthopsychiatry*, *84*(4), 364–376. https://doi.org/10.1037/h0099845

Huang, J., Li, Q., Xue, Y., Cheng, T., Xu, S., Jia, J., & Feng, L. (2015). TeenChat: A Chatterbot System for Sensing and Releasing Adolescents' Stress. *Health Information Science Lecture Notes in Computer Science*, 133–145. https://doi.org/10.1007/978-3-319-19156-0_14

Inkster, B., Sarda, S., & Subramanian, V. (2018). An Empathy-Driven, Conversational Artificial Intelligence Agent (Wysa) for Digital Mental Well-Being: Real-World Data Evaluation Mixed-Methods Study. *JMIR MHealth and UHealth*, *6*(11). https://doi.org/10.2196/12106

International Organization for Standardization. (2010). *Ergonomics of human-system interaction — Part 210: Human-centred design for interactive systems* (ISO Standard No.ISO 9241-210:2010):https://www.iso.org/standard/52075.html

Kessler, R. C., Angermeyer, M., Anthony, J. C., DE Graaf, R., Demyttenaere, K., Gasquet, I., DE Girolamo, G., Gluzman, S., Gureje, O., Haro, J. M., Kawakami, N., Karam, A., Levinson, D., Medina Mora, M. E., Oakley Browne, M. A., Posada-Villa, J., Stein, D. J., Adley Tsang, C. H., Aguilar-Gaxiola, S., Alonso, J., … Ustün, T. B. (2007). Lifetime prevalence and age-of-onset distributions of mental disorders in the World Health Organization's World Mental Health Survey Initiative. World psychiatry : official journal of the World Psychiatric Association (WPA), 6(3), 168–176.

Kindness, P., Masthoff, J., & Mellish, C. (2017). Designing emotional support messages tailored to stressors. International Journal of Human-Computer Studies, 97, 1–22. https://doi.org/10.1016/j.ijhcs.2016.07.010

Kretzschmar, K., Tyroll, H., Pavarini, G., Manzini, A., & Singh, I. (2019). Can Your Phone Be Your Therapist? Young People's Ethical Perspectives on the Use of Fully Automated Conversational

Agents (Chatbots) in Mental Health Support. Biomedical Informatics Insights, 11, 1178222619829083–1178222619829083. https://doi.org/10.1177/1178222619829083

Laugwitz, B., Held, T., & Schrepp, M. (2008). Construction and Evaluation of a User Experience Questionnaire. Lecture Notes in Computer Science, 63–76. https://doi.org/10.1007/978-3-540-89350-9_6

Lin, Y.-N. (2001). Taiwanese University Students' Conceptions of Counseling. *Journal of Contemporary Psychotherapy*, *31*(3), 200–211.

Matsumoto, D. (1989). Cultural Influences on the Perception of Emotion. Journal of Cross-Cultural Psychology, 20(1), 92-105. https://doi.org/10.1177/0022022189201006

Nielsen, J., & Landauer, T. (1993). A mathematical model of the finding of usability problems. Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems, 206–213. https://doi.org/10.1145/169059.169166

O'dea, B., Achilles, M. R., Werner-Seidler, A., Batterham, P. J., Calear, A. L., Perry, Y., … Christensen, H. (2018). Adolescents' Perspectives on a Mobile App for Relationships: Cross-Sectional Survey. *JMIR MHealth and UHealth*, *6*(3). https://doi.org/10.2196/mhealth.8831

Park, J.-J. (2018). A Development of Chatbot for Emotional Stress Recognition and Management using NLP. *The Transactions of the Korean Institute of Electrical Engineers*, *67*, 954–961. https://doi.org/http://doi.org/10.5370/KIEE.2018.67.7.954

Pearl, C. (2017). Designing voice user interfaces : principles of conversational experiences (First edition.). O'Reilly Media. https://learning.oreilly.com/library/view/designing-voice-user/9781491955406/

Sauro, J., & Lewis, J. R. (2016). Quantifying the user experience: practical statistics for user research. ProQuest Ebook Central. Morgan Kaufmann. https://ebookcentral.proquest.com/lib/tampere/reader.action?docID=4592083

Schrepp, Martin. (2015). User Experience Questionnaire Handbook. https://doi.org/10.13140/RG.2.1.2815.0245

TWNIC. (December 31, 2018). Leading communication software among online messaging users in Taiwan as of November 2018 [Graph]. In Statista. Retrieved December 29, 2020, from https://www-statista-com.libproxy.tuni.fi/statistics/968238/taiwan-communication-software-usage/

Wang, C.-chi, & Hsaio, W. (2007). Life Stress, Depressive Experience and Resilience of Junior High School Students. Journal of Living Sciences, 11, 1–31.

Wang, Y., Chai, J., &; Lei, T. (2021). Chinese UMUX: Cross-Cultural Adaptation for Perceived Usability Measurement. International Journal of Human–Computer Interaction, 1–13. https://doi.org/10.1080/10447318.2021.1898826

Wu, C., Liu, S., Chang, S., & Sun, F. (2013). Surveys of medical seeking preference, mental health literacy, and attitudes toward mental illness in Taiwan, 1990–2000. Journal of the Formosan Medical Association, 113(1), 33–41. https://doi.org/10.1016/j.jfma.2013.03.004

Zgambo, M., Kalembo, F., Guoping, H., & Honghong, W. (2012). Depression Among Chinese Children And Adolescents: A Review Of The Literature. *International Journal of Child, Youth and Family Studies*, *3*(4.1), 442. https://doi.org/10.18357/ijcyfs34.1201211543

# Appendix A

Interview plan that was drafted in English first, relevant parts were translated into Chinese afterwards for the actual interviews.

## Conducting Research

**I'm looking to answer these questions:**

1. What are the particular needs of Taiwanese Adolescents when discussing stress with experts/school councilors?
2. How do experts/school counselors help Taiwanese adolescents in relieving stress
    a. What information do schools provide to support them?
3. What are the most common issues school counselors face when trying to help Taiwanese adolescents relieve stress?

**Style**

Interview questions will be very contextual, but overall will try to answer the above and try to follow the below flow.

**Flow**

Counselors data > what are the stressors students have > how are they handled by counselors> what are difficulties the students have in utilizing the support given > what methods do the counselors use to solve this issue > Does the school provide any support or outreach promotion in giving students access to mental support? > Are there any efforts to digitize some of the processes?

**Example and question database**

- 你可以先跟跟我講一下你自己跟你做輔導老師做多久嗎？
  Can you first tell me a bit about yourself and how long you have been working as a councilor?
- 輔導學生的工作主要是在做什麼
  What does your work entail?

- 你在輔導時，主要輔導學生什麼樣的問題？
  What problems do you have to help solve for students?
- 你是怎們輔導他們的
  How do you help them?
- 在輔導時，會不會遇到一些特別困難的問題？是什麼樣的問題？你怎麼解決的？
  What, if any, are the particular difficulties you can face in the counseling sessions and how do you address that?

- 就你的論點來看，台灣的學生是否有比較大的壓力？
  Do you think in general that students have too much stress in Taiwan?
- 為什麼這麼認為？
  Why?
- 你覺得台灣的學生是否有缺法舒壓的認知？
  What do you think is lacking for the students when they have stress and can't relieve it?

- 為什麼這麼認為？
  Why do you think so?

- 台灣學生是否有困難表達打或溝通他們的問題跟情緒？
  Do you think Taiwanese students have trouble expressing their emotions and problems?
- 你是怎麼協助他們表達問題呢？
  How do you help them with expressing their problems and stress situations?

- 那本校是怎麼去協助學生舒壓呢？
  What are the school's official efforts in helping their students cope with stress?
- 你認為學生是否對學校的宣傳有足夠的認知?
  Do you think the students are aware of the school's efforts?

- 你認為學生他們目前藉由什麼方式去舒壓？
  What do you think students rely most on for stress relief? (family, friends, self improvement?)
- 有任何學生會自我尋求紓壓技巧嗎？像是自我處理技巧之類的？
  Do you have any experience with students seeking self help methods for stress relief?

- 在輔導時，你有嘗試過給學生這些紓壓技巧嗎？
  During counseling, have you ever provided some techniques that can help them relive stress?

- 學生是什麼樣反應？
  What were the students' reaction and comments on this?

**Opening**
你好！我叫鄭力嘉，謝謝你同意接受專訪。今天我想跟你聊聊有關你的輔導工作。這主要是想透過跟輔導老師的轉訪，了解你們在替學生輔導時，主要是做什麼。接到學生第一次輔導的要求時，怎麼去讓它願意開啟心胸跟你分享問題，以及怎麼去鼓勵學生分享他們的問題，好幫助他們舒壓。如果在訪談中有任何不方便分享的事情的話，我隨可以直接跳過。我會盡量控制訪談時間在1小時內。如果沒問題的話，我們開始吧。

Hi (Name), thanks for agreeing to take part in my research interview. My name is James and today I want to talk to you and understand a bit about your occupation as a school counselor. It's not a test, I'm mainly looking to hear what you usually do when you first get a counseling case, when you start your counseling session and how you go about helping the student in releasing their stress. If at any time you don't feel like answering a question, let me know and I'll stop. This interview is not going to be very long, I'll keep it in the hour mark. Shall we begin?

**Closing**

真的很謝謝老師您的協助！差不多就這樣！我學到了很多有助於我的研究的知識。你有什麼想加的或想問的嗎？那我可以們再保持聯絡嗎，有後續會通知你的。謝謝你！
Thanks for your participation! I think that's about it and what I've learned from you has been extremely helpful. Do you have any comments for me in general or anything you would like to add? Would you like to stay updated on this? Once again thanks for the help!

## Appendix B

Below is the user scenario that was used for the user testing. Before the participants interacted with the prototype, the scenario was communicated to them verbally. Each task was told to the participants only after they completed the prior task before.

**Test Scenario:**
你是國中三年級生。最近因大會考即將來臨，學業壓力變得很重，開始影響到生活作息。
在校園的海報上看到學校有新的服務能幫助你紓壓。
你並掃了QR碼，用line加了小煩聊天機器人。

You are currently a 9th grade middle school student.
Due to the upcoming CAP exam, pressure to perform well academically has been mounting, which has led to you being stressed in your everyday life. At school, you noticed a poster mentioning that a new service is available to help you relieve some stress. You scan the QR code with your phone and add the Chatbot as a friend.

**Task 1.**
請跟小凡開始聊天，並了解他怎麼協助你
Please start chatting with the chatbot and learn how it can help you.

**Task 2.**
請跟小凡用聊天的方式，舒壓
Please use the chatting function to relieve some stress

**Task 3.**
請跟小凡學習紓壓技巧
Please use the technique fuction to learn some methods to deal with stress from the chatbot

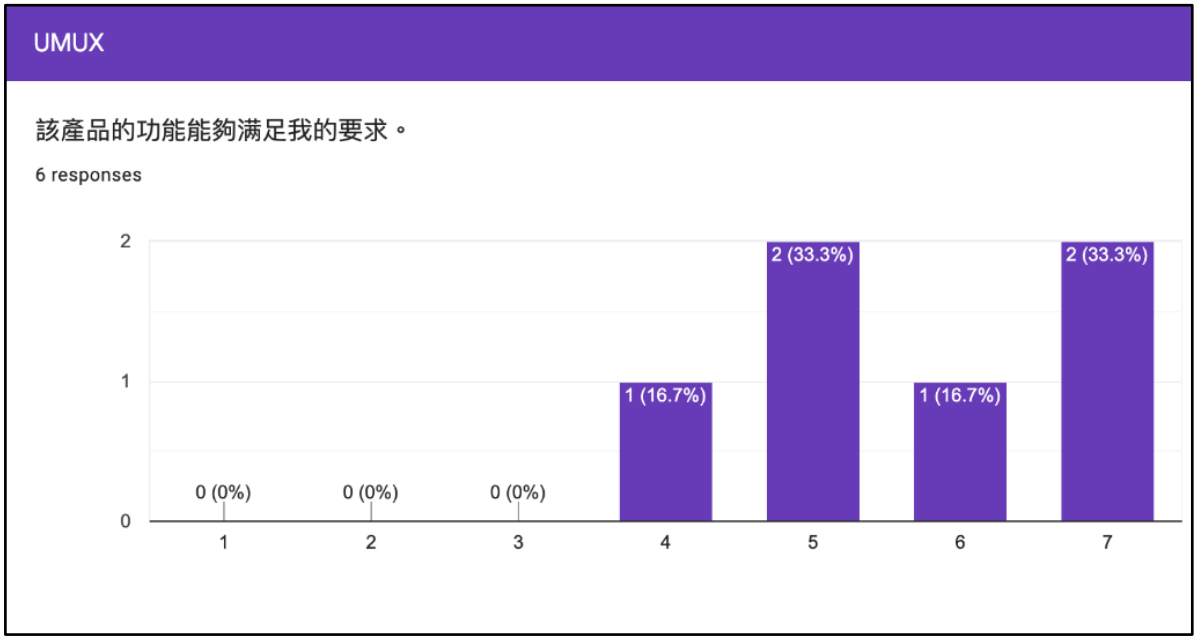
**Task 4.**
請跟小凡詢問更能幫助的你真的人
Please enquire with the chatbot about real people that can assist further.

## Appendix C

### Usability Metric User Experience (UMUX)

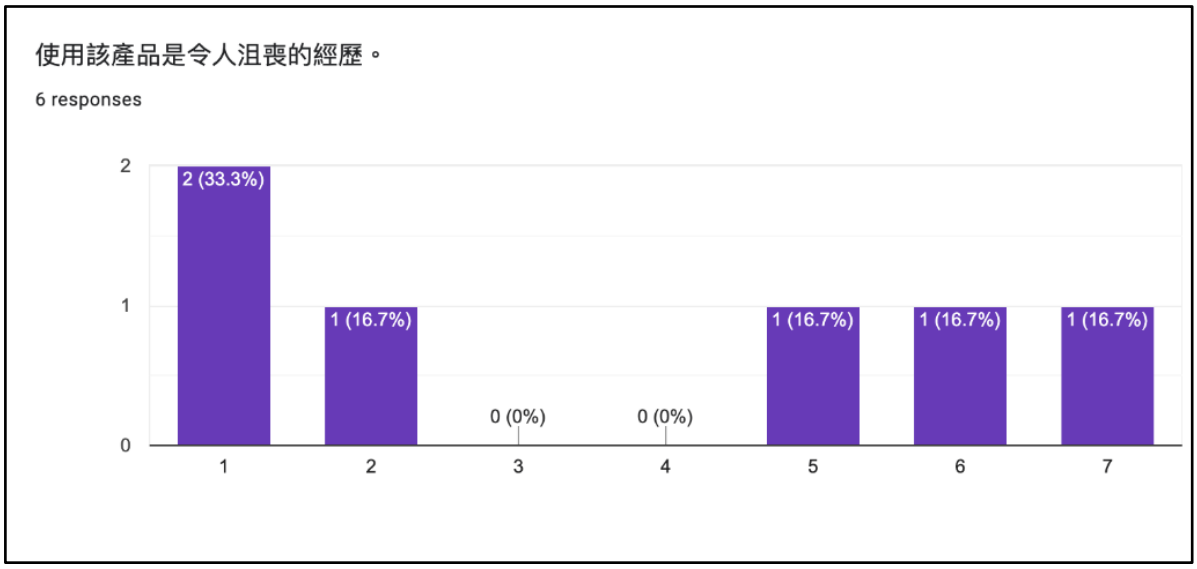
Questions 1 through 4 are based on a cross culturally adapted version of the Usability Metric User Experience (Wang et al., 2021), modified from Simplified Mandarin Chinese to Traditional Mandarin Chinese.

*Question 1: This product's capabilities meet my requirements.*



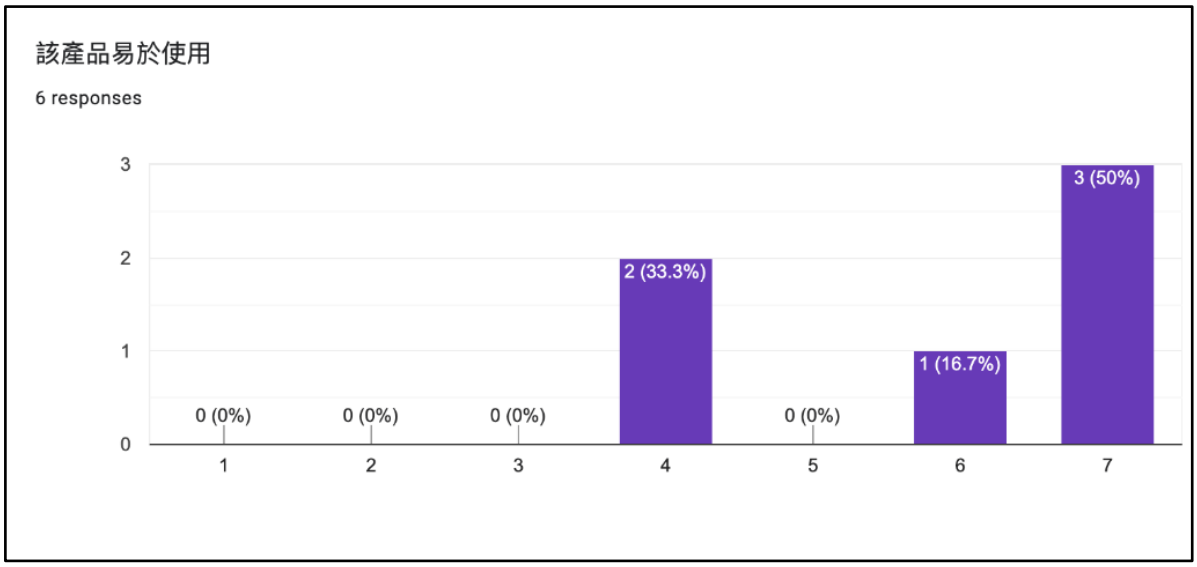

*Note*. 1: Strongly Disagree ~ 7:Strongly Agree

*Question 2: Using this product is a frustrating experience.*



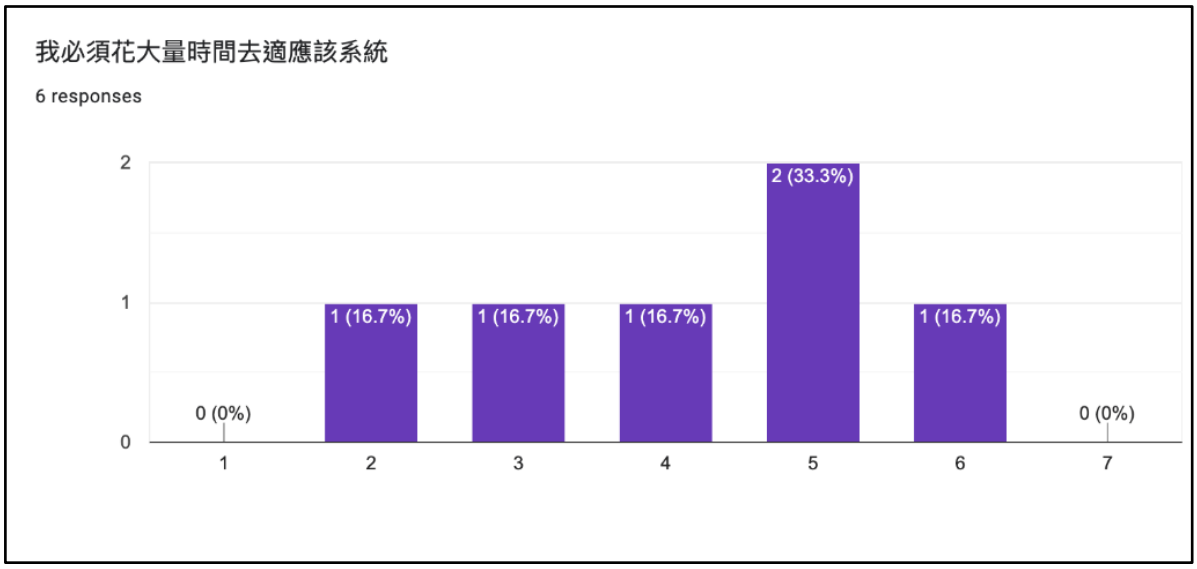

*Note.*1: Strongly Disagree ~ 7:Strongly Agree

*Question 3: This product is easy to use.*




*Note.* 1: Strongly Disagree ~ 7:Strongly Agree

*Question 4: I have to spend too much time correcting things with this product.*

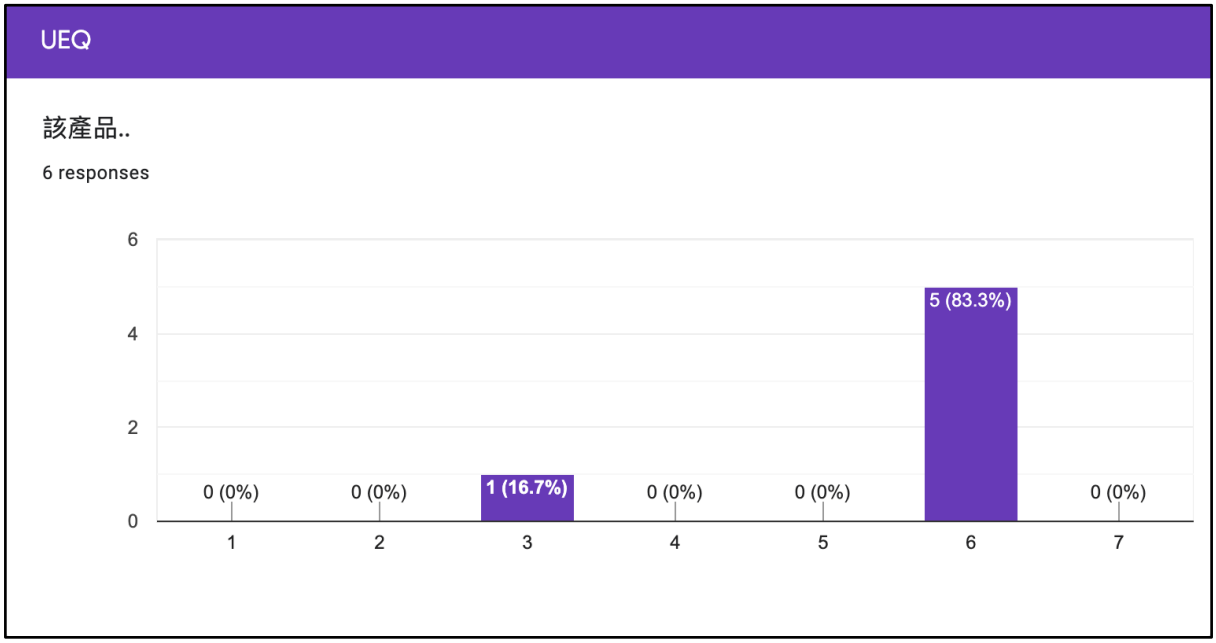



*Note.* 1: Strongly Disagree ~ 7:Strongly Agree

**User Experience Questionnaire(UEQ)**

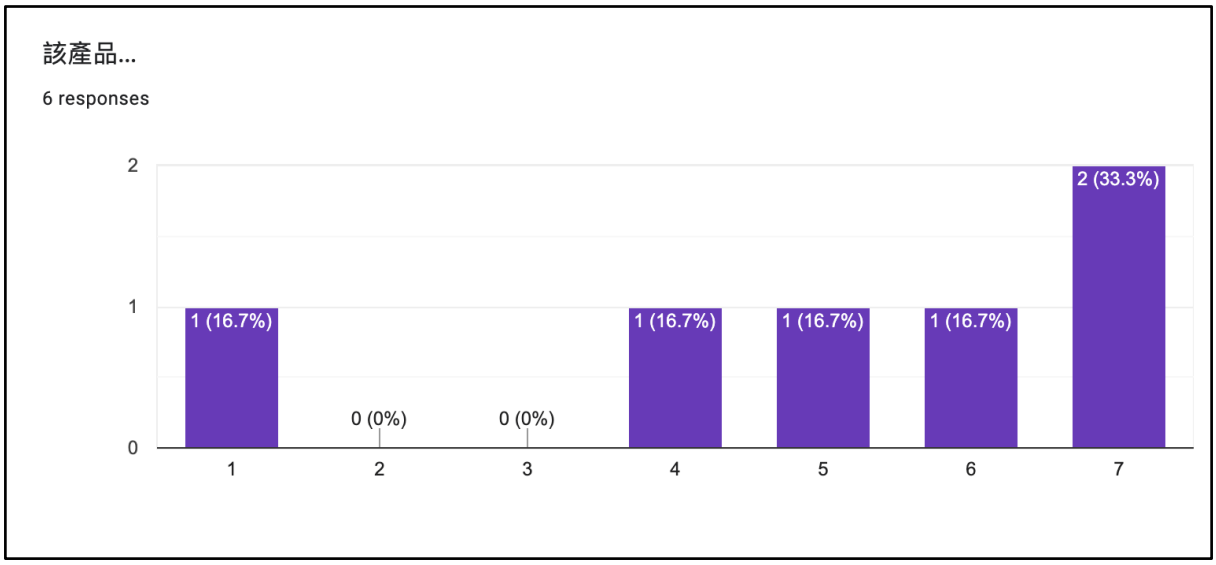
Questions 5 through 30 are based on the Simplified Mandarin Chinese version of the User Experience Questionnaire(Laugwitz et al., 2008), modified to Traditional Mandarin Chinese. The Simplified Mandarin Chinese can be found here: https://www.ueq-online.org/

*Question 5: This product is...*



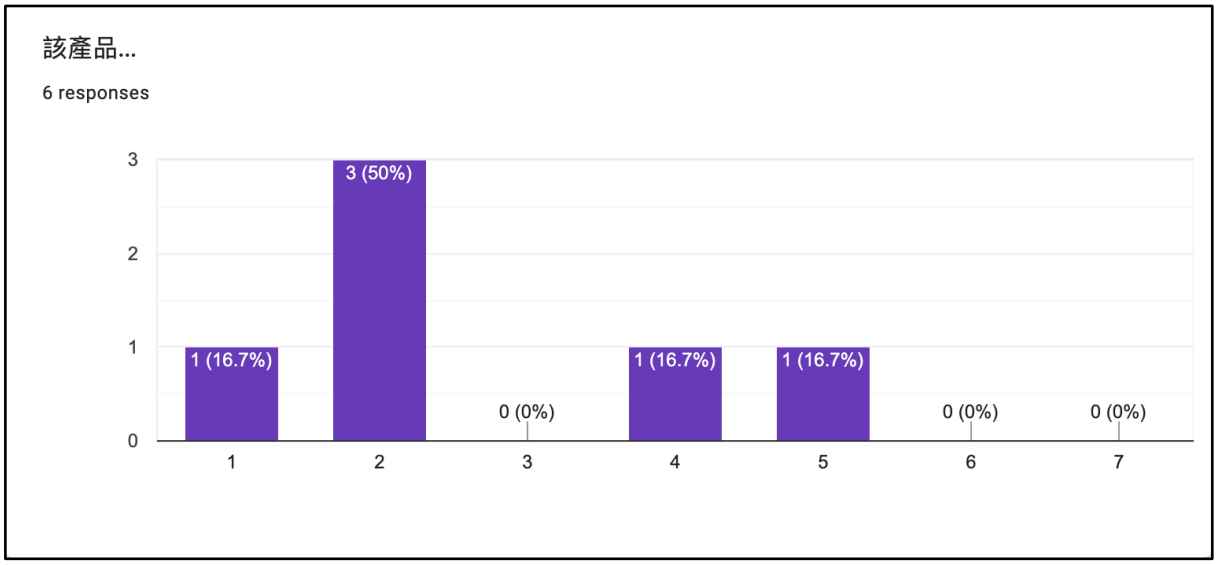

*Note.* 1:annoying ~ 7:enjoyable

*Question 6: This product is...*



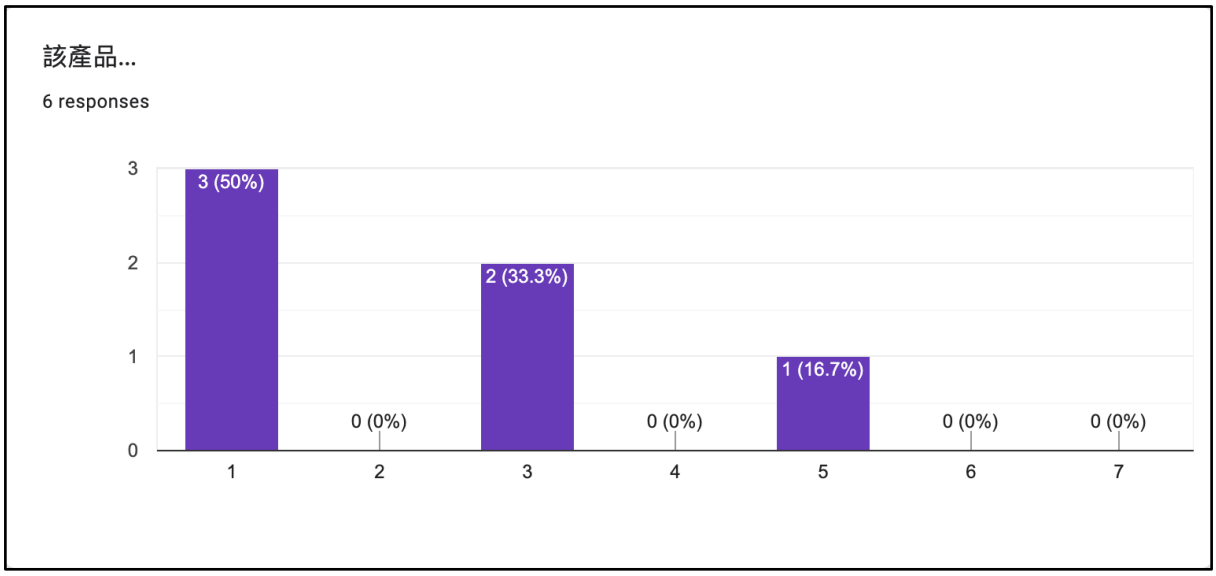

*Note.* 1:not understandable ~ 7:understandable

*Question 7: This product is...*



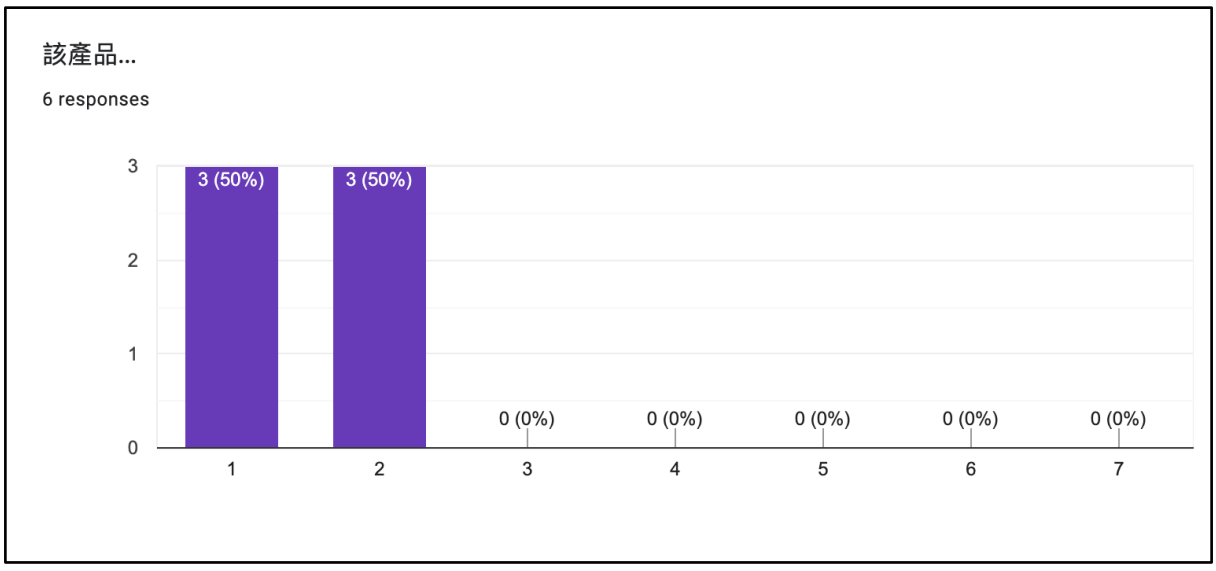

*Note.* 1:creative ~ 7:dull

*Question 8: This product is...*



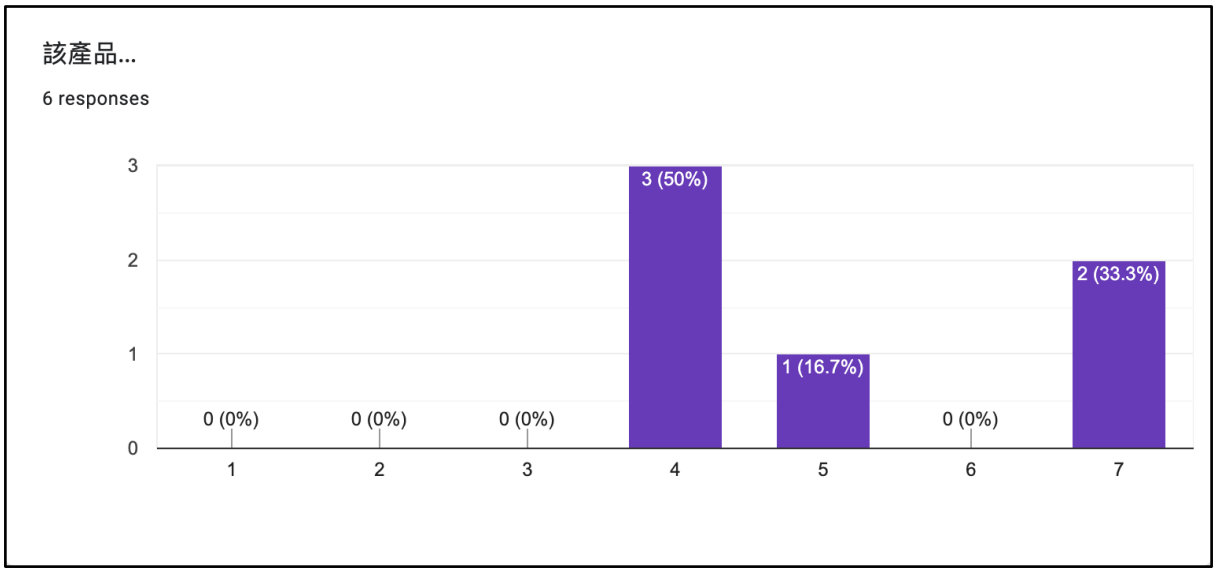

*Note.* 1:easy to learn ~ 7:difficult to learn

*Question 9: This product is...*



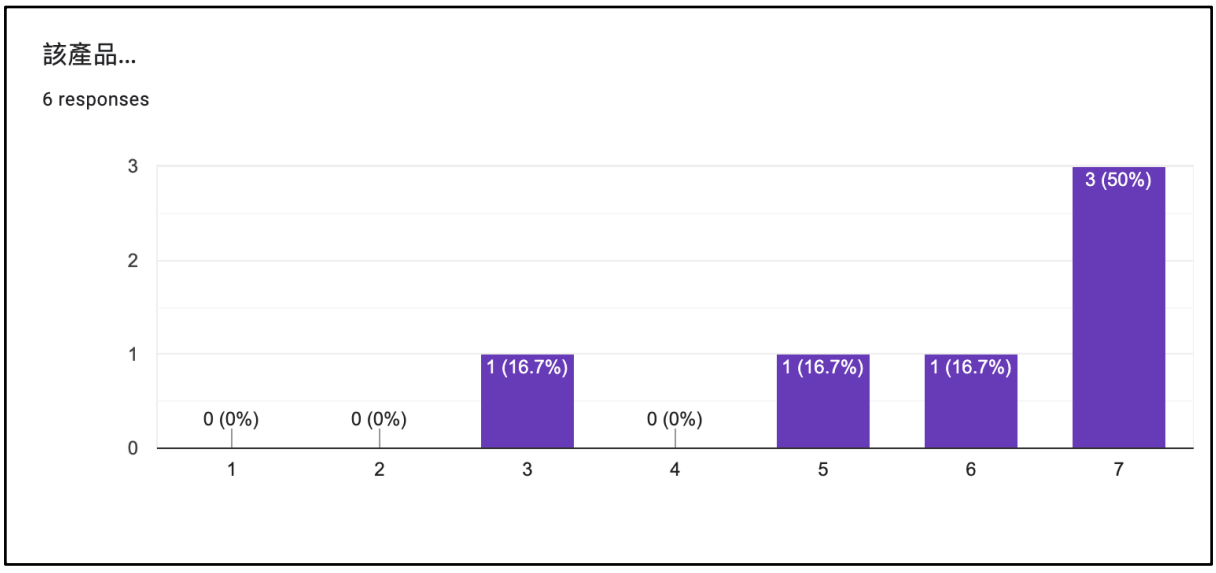

*Note.* 1:valuable ~ 7:inferior

*Question 10: This product is...*



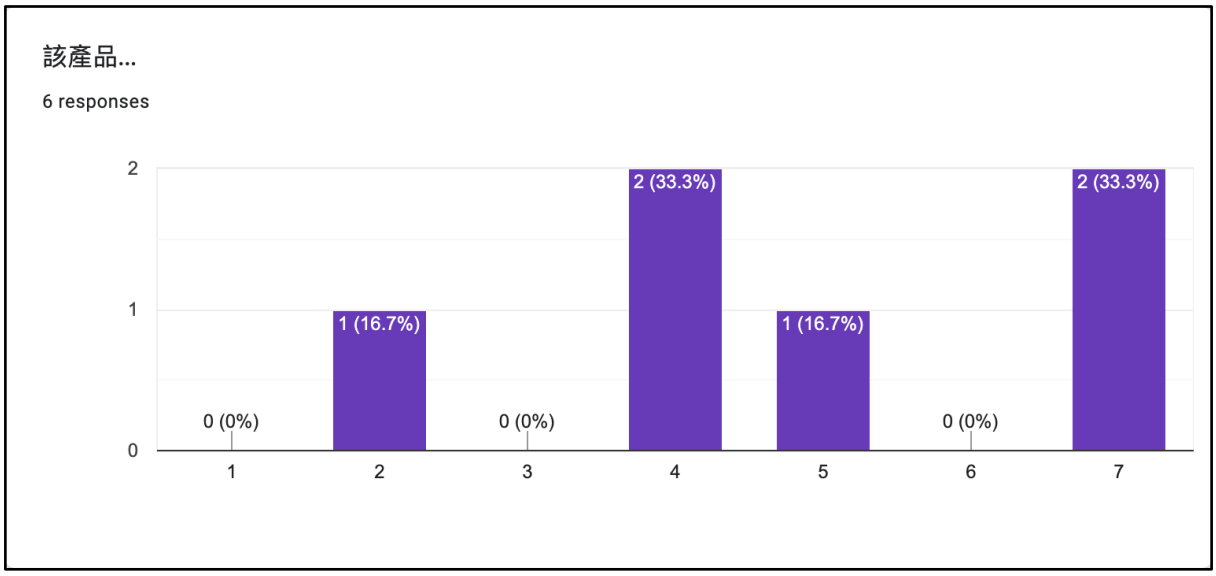

*Note.* 1:boring ~ 7:exciting

*Question 11: This product is...*



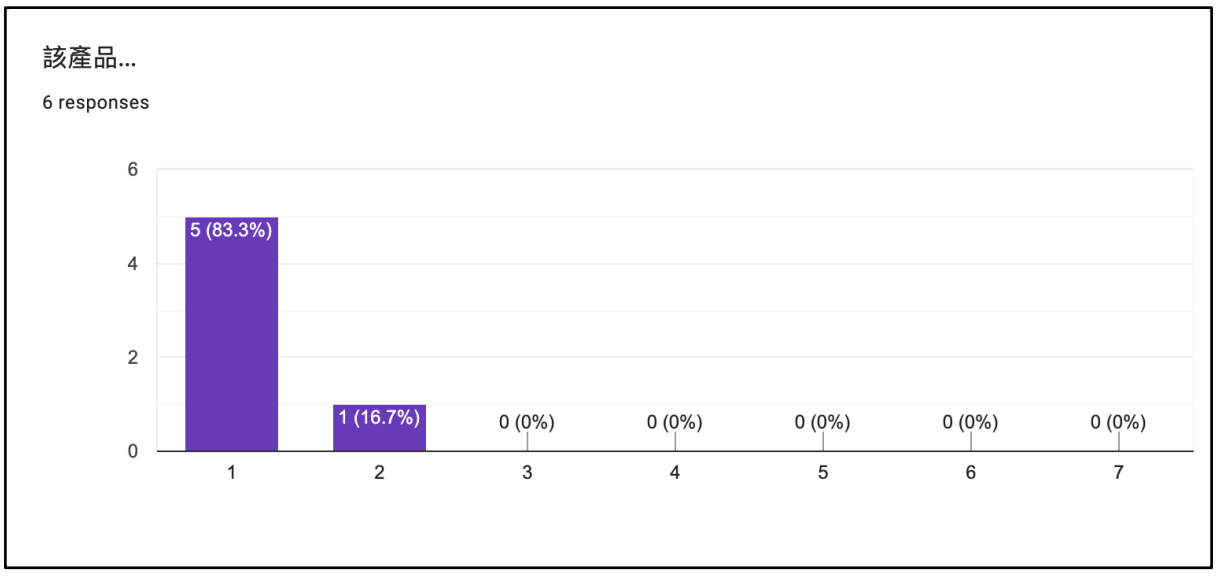

*Note.* 1:not interesting ~ 7:interesting

*Question 12: This product is...*



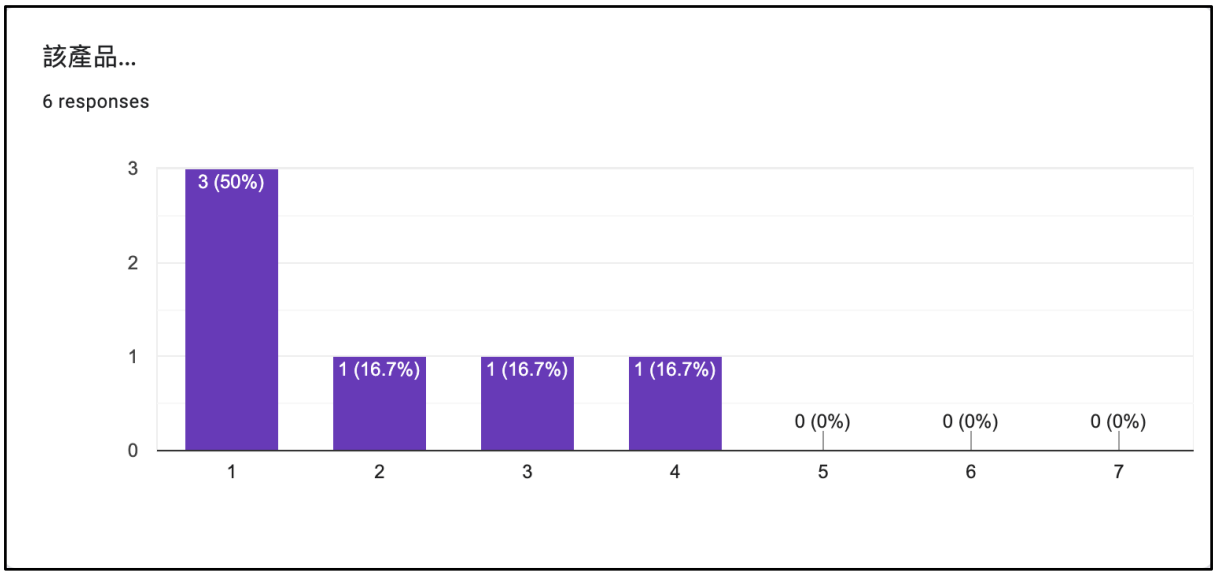

*Note.* 1:unpredictable ~ 7:predictable

*Question 13: This product is...*



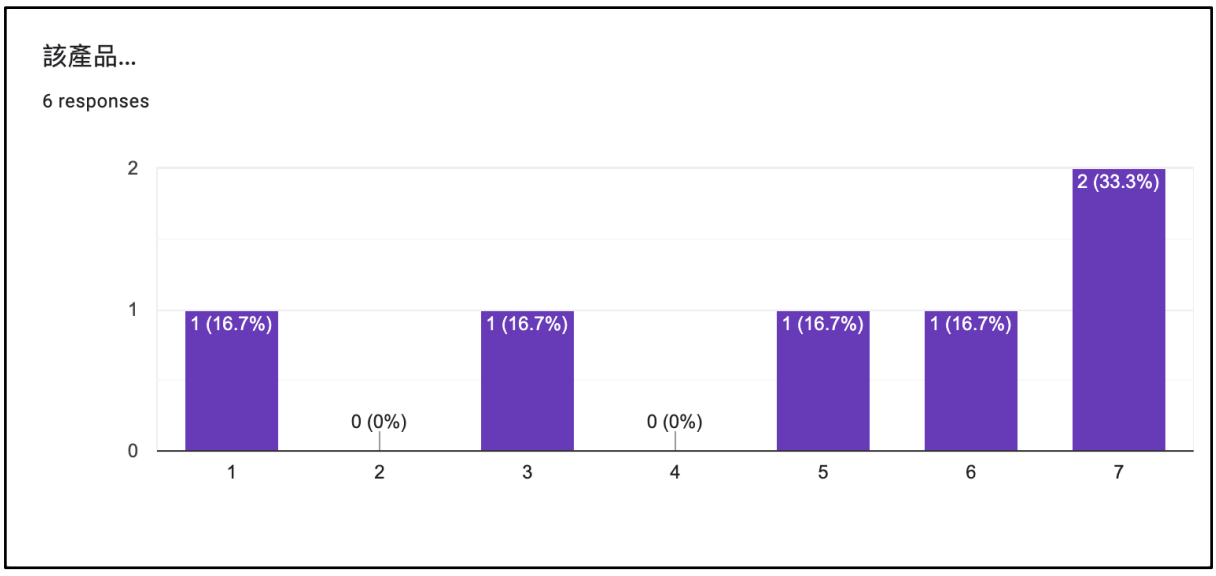

*Note.* 1:fast ~ 7:slow

*Question 14: This product is...*



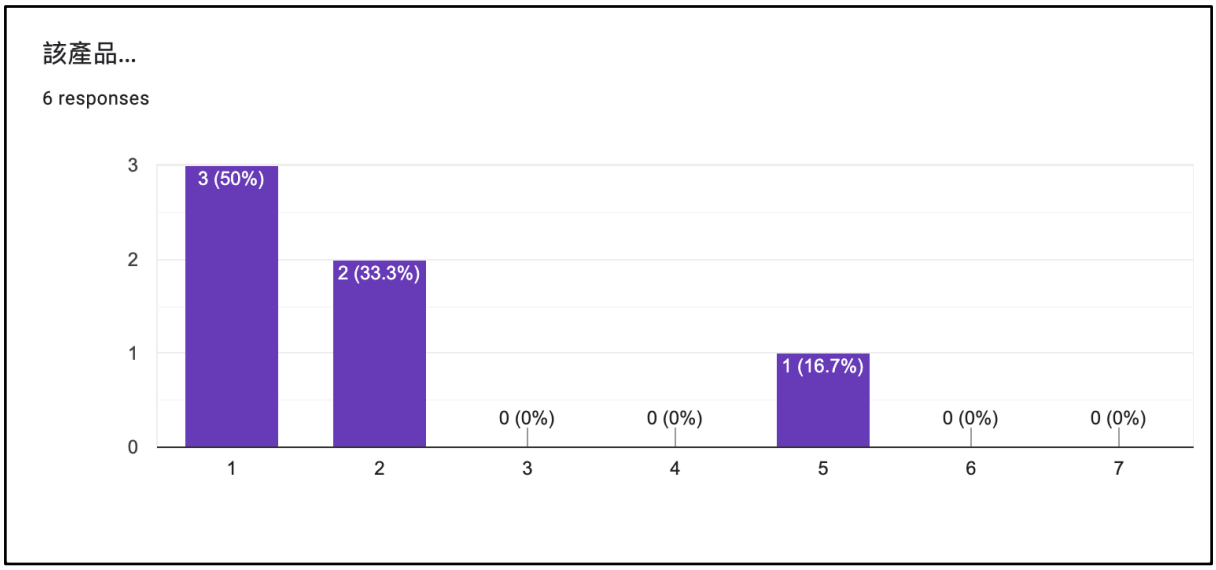

*Note.* 1:inventive ~ 7:conventional

*Question 15: This product is...*



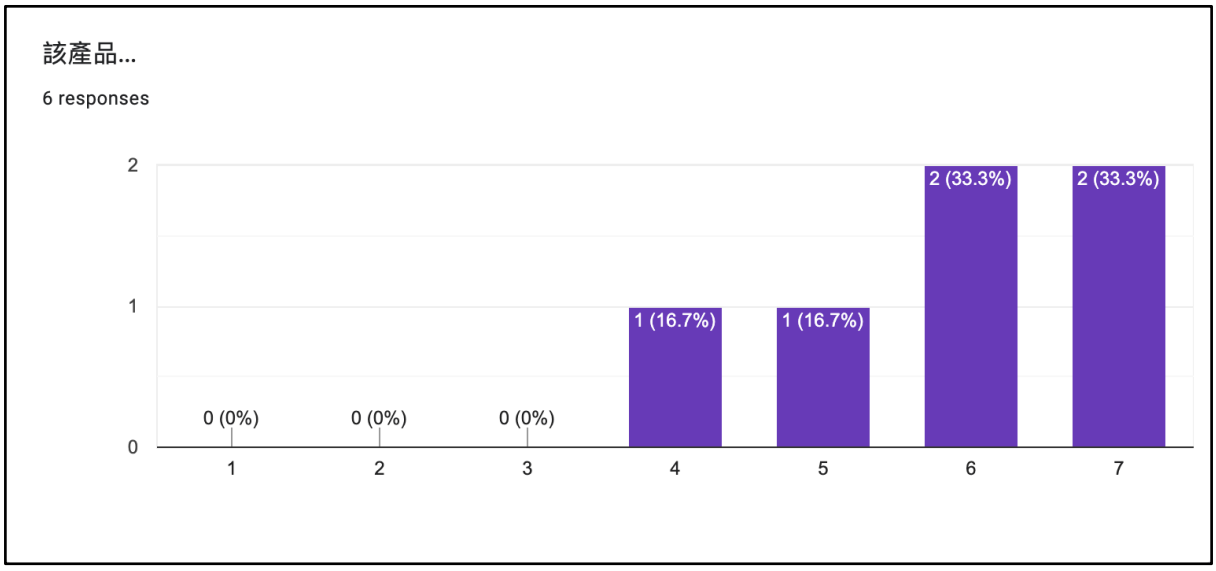

*Note.* 1:obstructive ~ 7:supportive

*Question 16: This product is...*



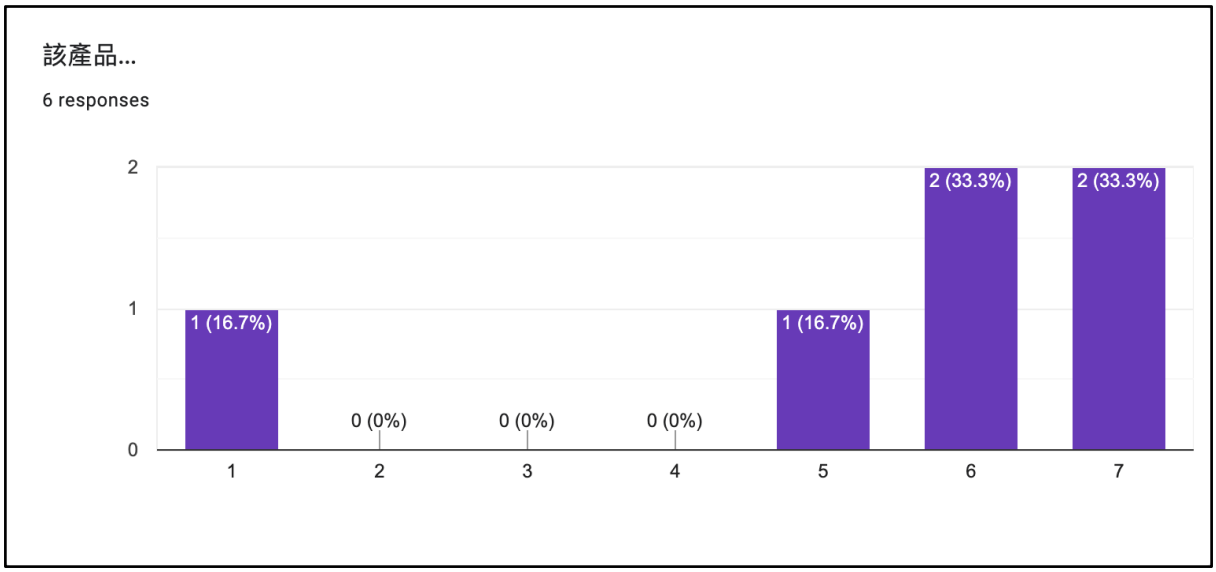

*Note.* 1:good ~ 7:bad

*Question 17: This product is...*



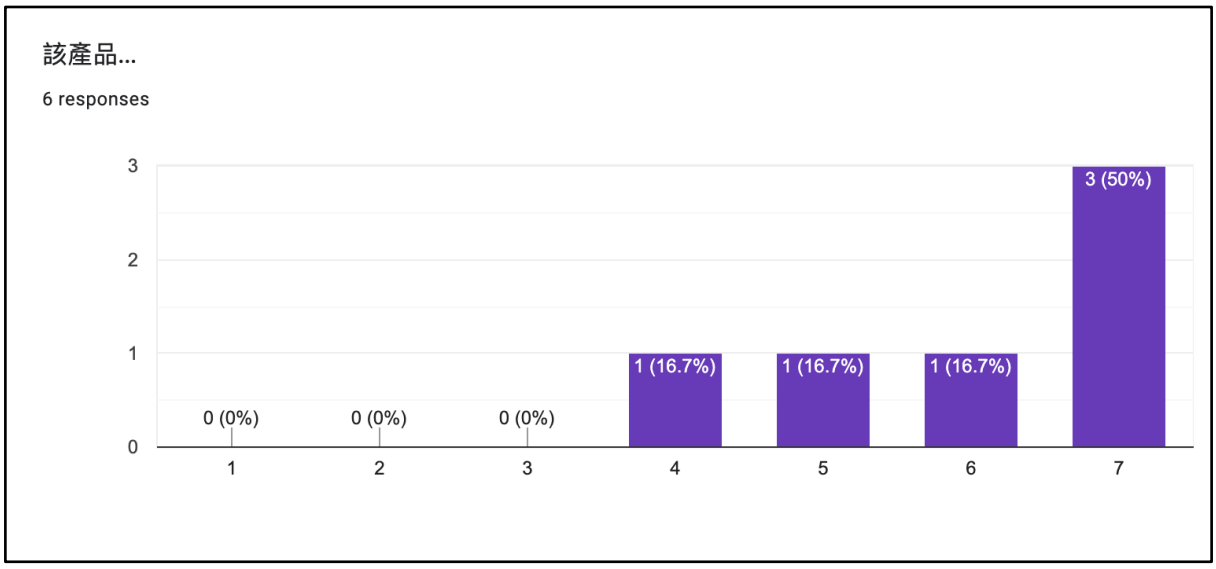

*Note.* 1:complicated ~ 7:easy

*Question 18: This product is...*



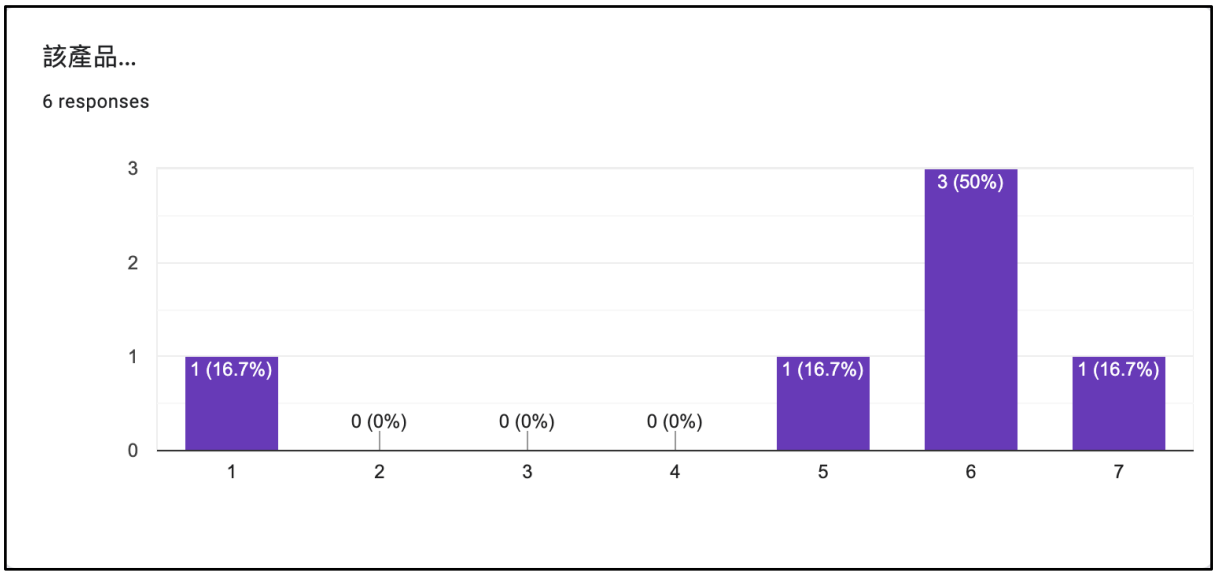

*Note.* 1:unlikable ~ 7:pleasing

*Question 19: This product is...*



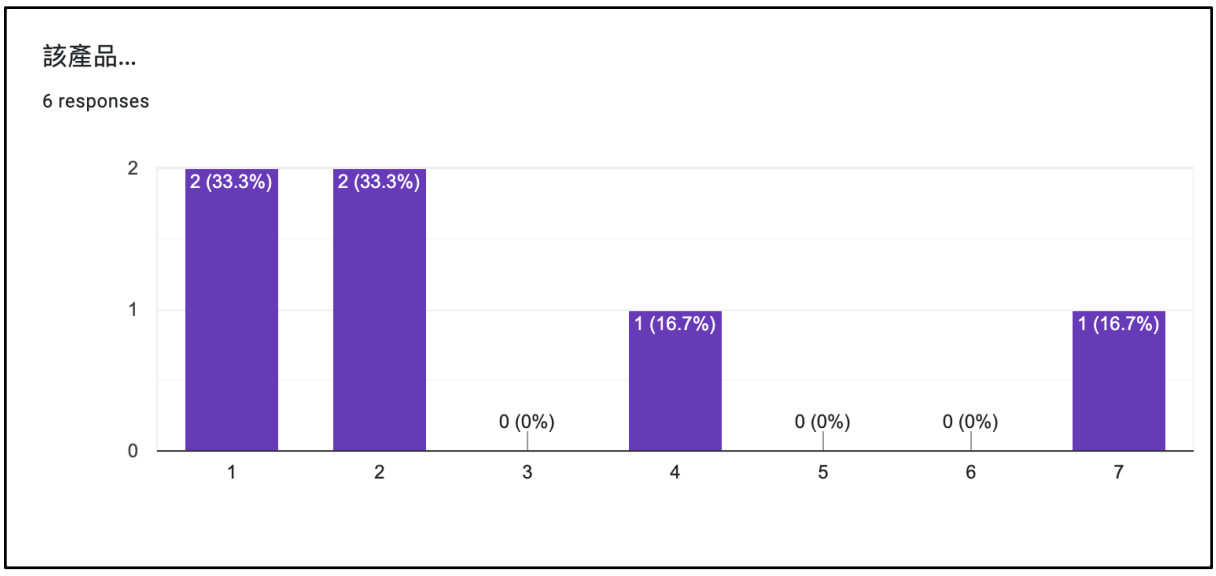

*Note.* 1:usual ~ 7:leading edge

*Question 20: This product is...*



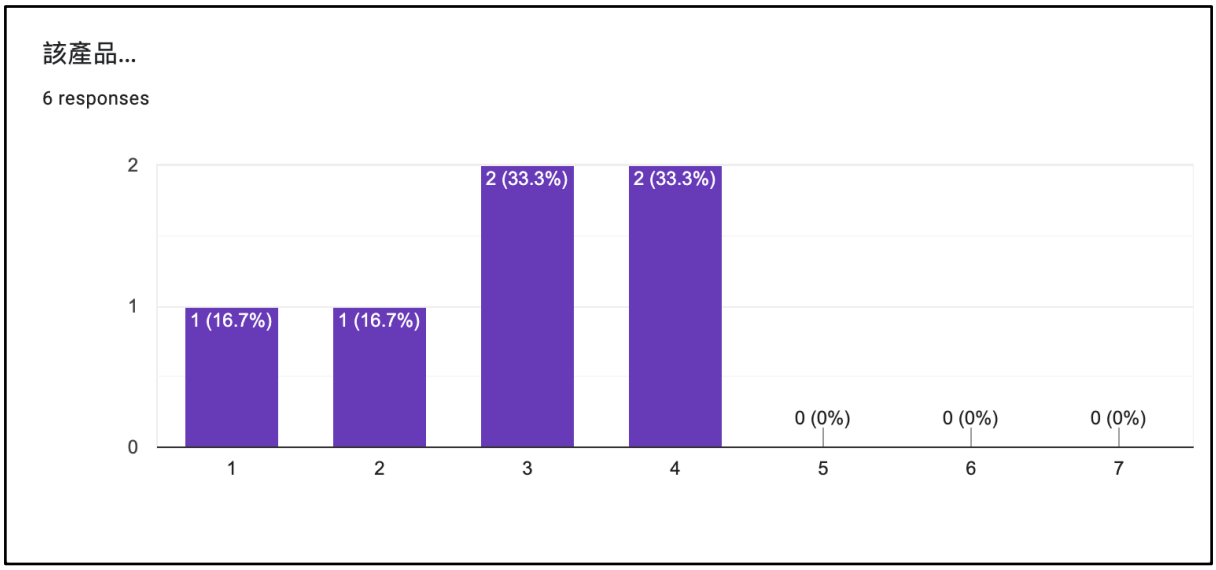

*Note.* 1:unpleasant ~ 7:pleasant

*Question 21: This product is...*



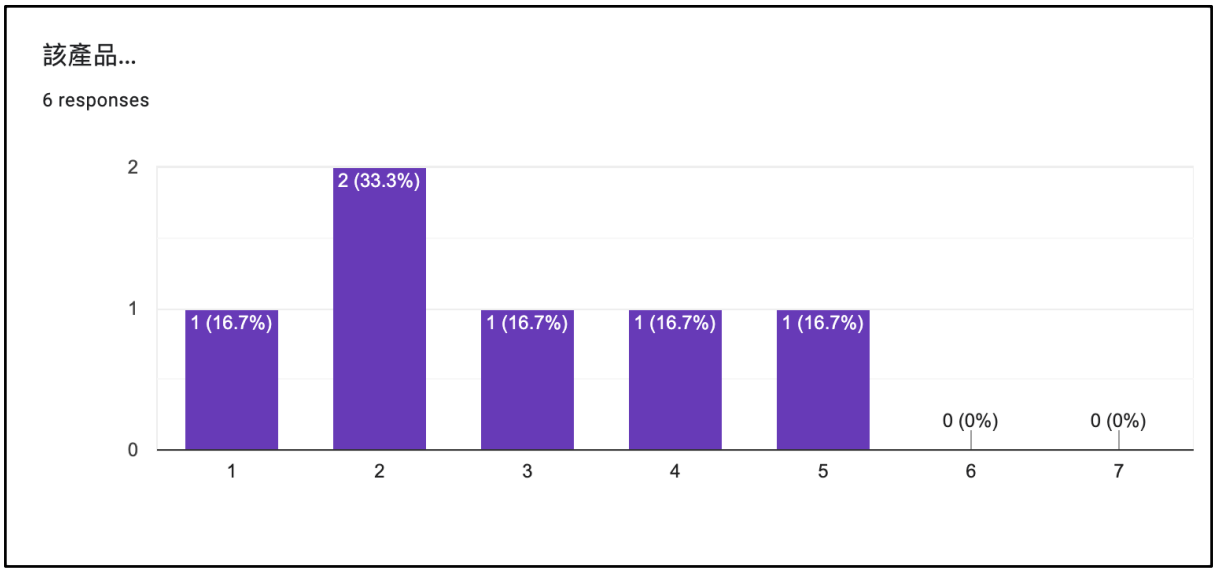

*Note.* 1:secure ~ 7:not secure

*Question 22: This product is...*



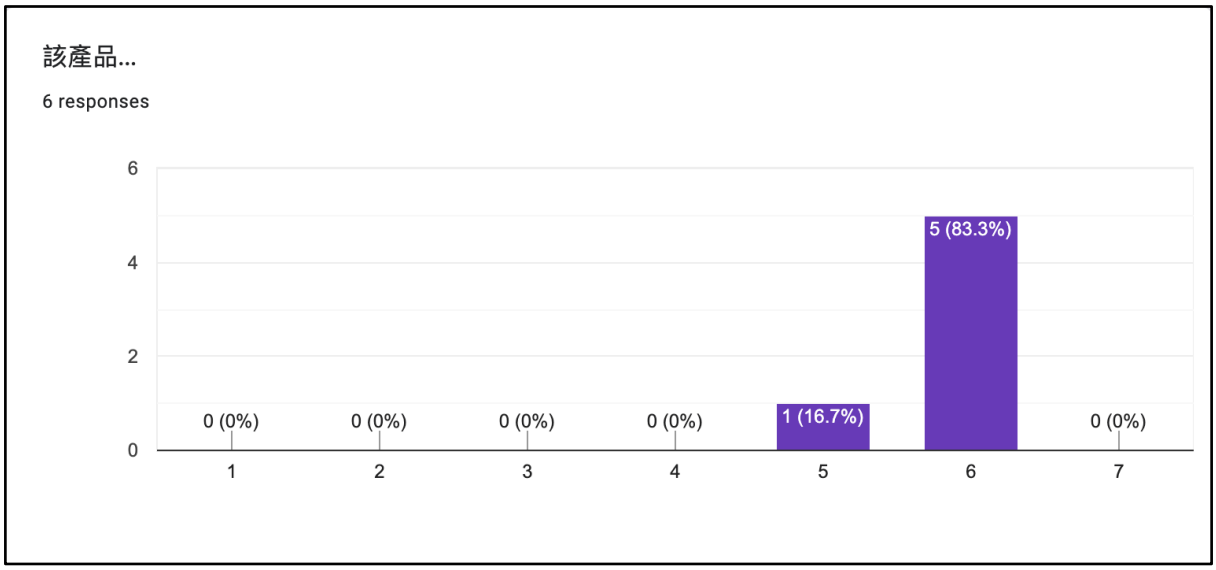

*Note.* 1:motivating ~ 7:demotivating

*Question 23: This product is...*



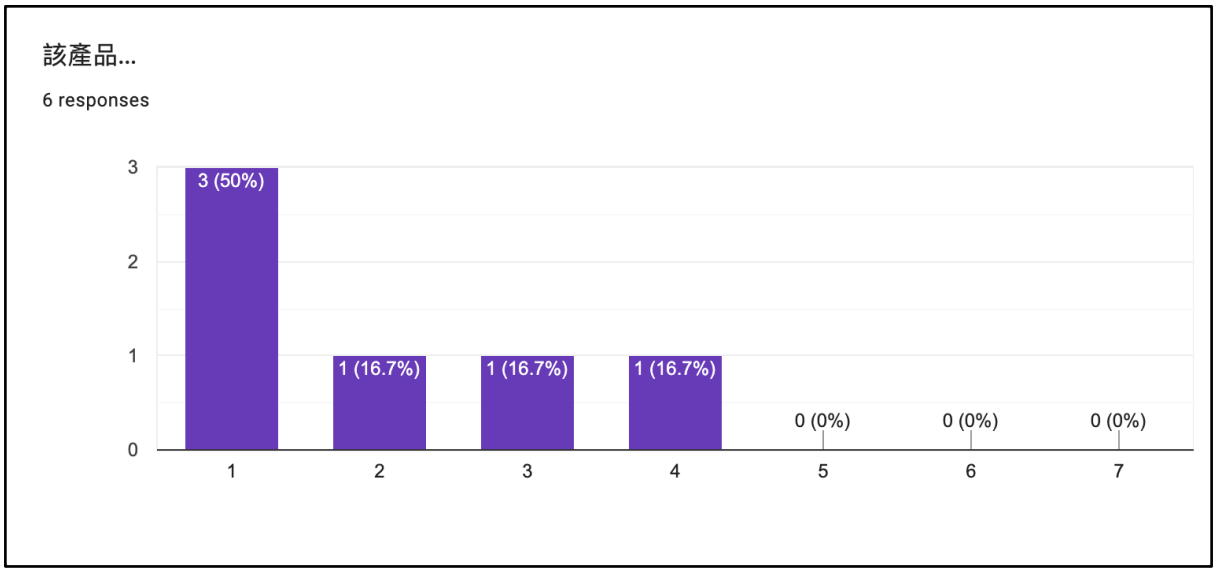

*Note.* 1:meets expectations ~ 7:does not meet expectations

*Question 24: This product is...*



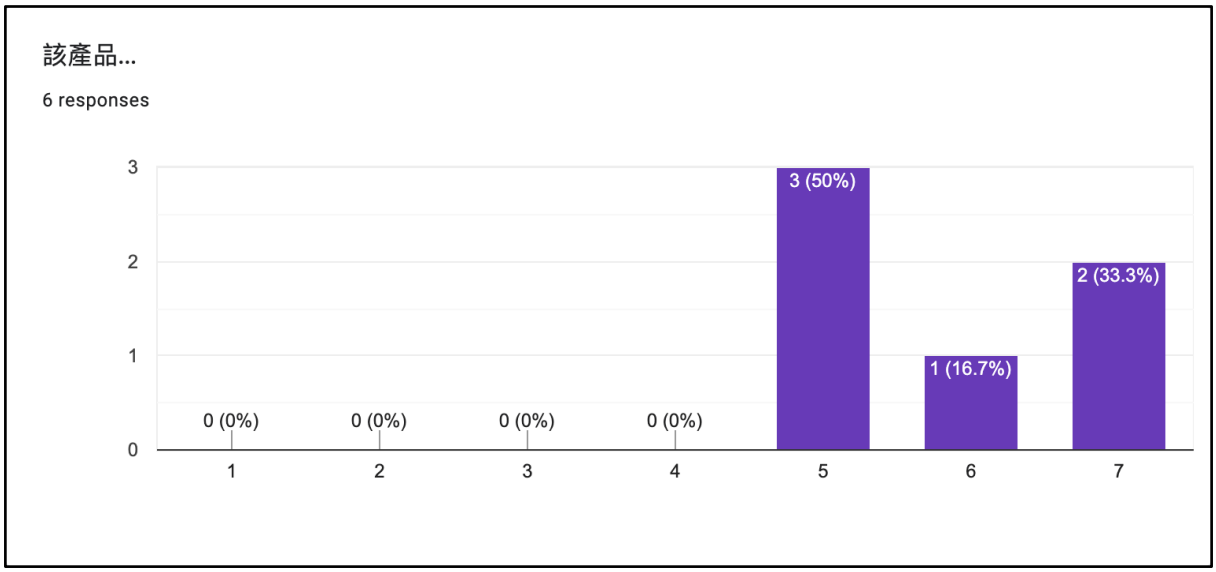

*Note.* 1:inefficient ~ 7:efficient

*Question 25: This product is...*



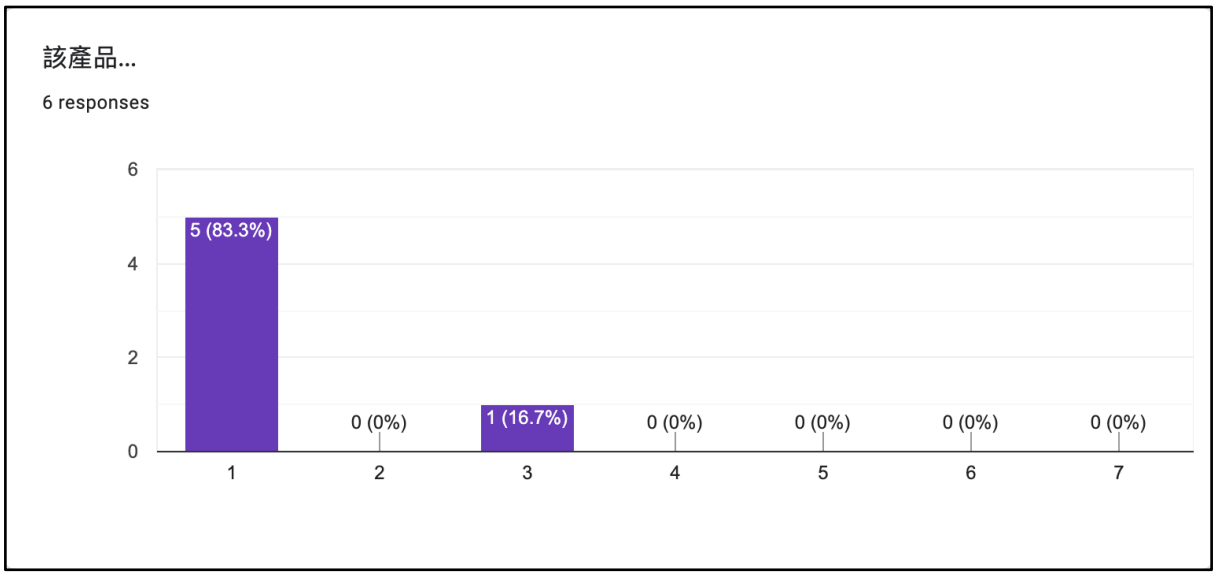

*Note.* 1:clear ~ 7:confusing

*Question 26: This product is...*



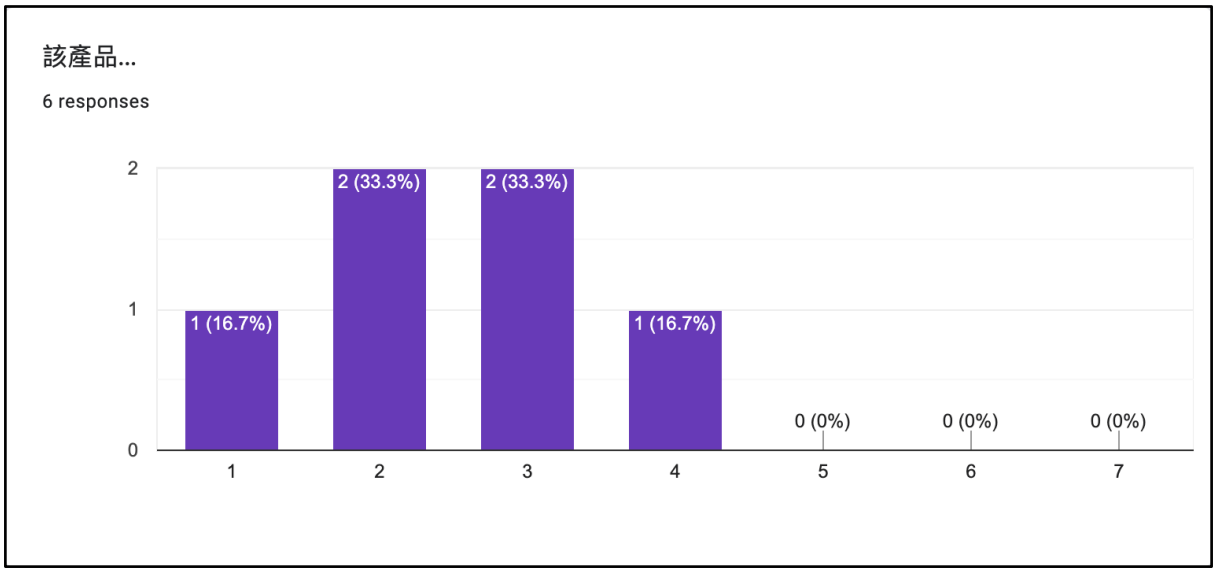

*Note.* 1:impractical ~ 7:practical

*Question 27: This product is...*



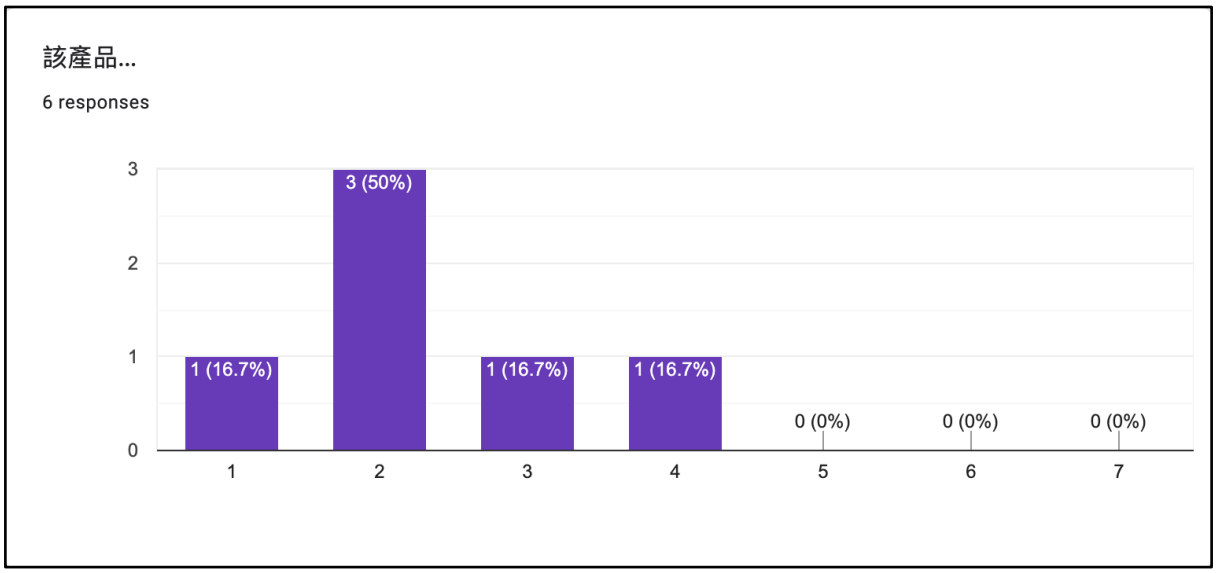

*Note.* 1:organized ~ 7:cluttered

*Question 28: This product is...*



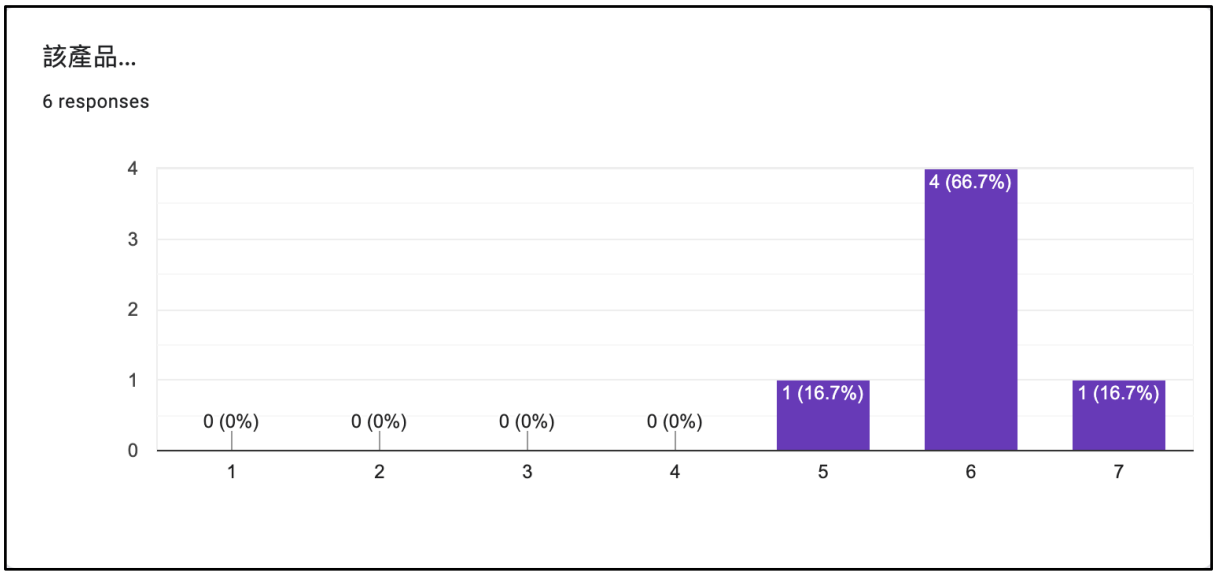

*Note.* 1:attractive ~ 7:unattractive

*Question 29: This product is...*




*Note.* 1:friendly ~ 7:unfriendly

*Question 30: This product is...*




*Note.* 1:conservative ~ 7:innovative

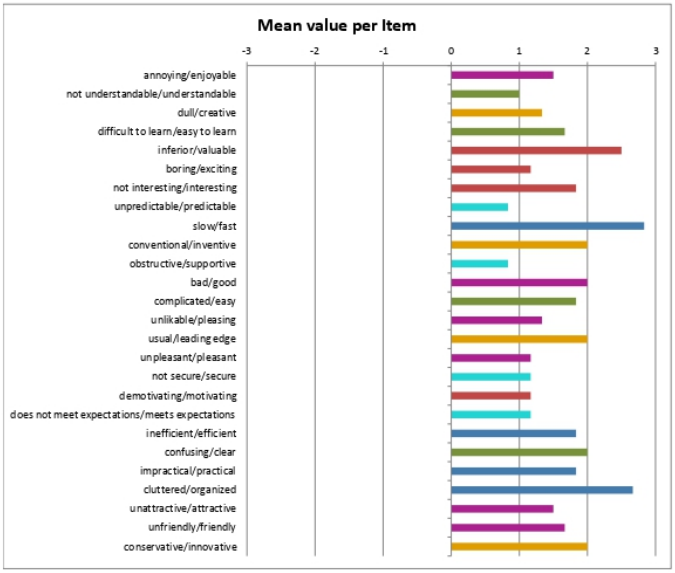
# Appendix D

This is an excerpt of the results page data calculated using the data analysis tool found on UEQ's website here: https://www.ueq-online.org/

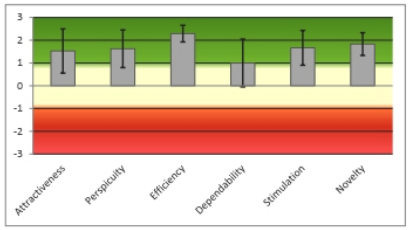
**Results**

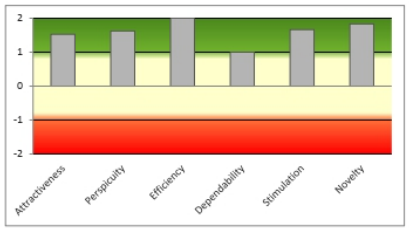
You can interpret the means of the scales. The UEQ does not produce an overall score for the user experience (you need to use the KPI extension for this). Because of the construction of the questionnaire it does make no sense to built such an overall score (for example by calculating the mean over all scales), since this value can not be interpreted properly. The values for the single items are listed to allow you to detect outliers in the evaluations. If an item shows big deviations to the evaluations of the other items of the same scale this can be a hint that the item is misinterpreted (for example, because of a special context in your evaluation) by a higher number of participants.
Values between -0.8and 0.8 represent a more or less *neutral evaluation* of the corresponding scale, values > 0,8 represent a *positive evaluation* and values < -0,8 represent a *negative evaluation* .
The range of the scales is between -3 (horribly bad) and +3 (extremely good). But in real applications, in general, only values in a restricted range will be observed. It is due to the calculation of means over a range of different persons with different opinions and answer tendencies (for example the avoidance of extreme answer categories) extremely unlikely to observe values above +2 or below -2.
Thus, even a quite good value of +1.5 for a scale looks from the purely visual standpoint on a scale range of -3 to +3 not as positive as it really is. For this reason this sheet contains two variants for the figure that depicts the scale means. Use the figure with the reduced scale -2 to +2 if you communicate the results to persons that have not much knolwedge on the interpretation of this type of data and in situations where you don't want to explain in detail how building mean values and answer tendencies influence the observed data.

| Item | Mean | Variance | Std. Dev. | No. | Left | Right | Scale |
|---|---|---|---|---|---|---|---|
| 1 | 1,5 | 1,5 | 1,2 | 6 | annoying | enjoyable | Attractiveness |
| 2 | 1,0 | 5,2 | 2,3 | 6 | not understandable | understandable | Perspicuity |
| 3 | 1,3 | 2,3 | 1,5 | 6 | creative | dull | Novelty |
| 4 | 1,7 | 2,7 | 1,6 | 6 | easy to learn | difficult to learn | Perspicuity |
| 5 | 2,5 | 0,3 | 0,5 | 6 | valuable | inferior | Stimulation |
| 6 | 1,2 | 2,2 | 1,5 | 6 | boring | exciting | Stimulation |
| 7 | 1,8 | 2,6 | 1,6 | 6 | not interesting | interesting | Stimulation |
| 8 | 0,8 | 3,8 | 1,9 | 6 | unpredictable | predictable | Dependability |
| 9 | 2,8 | 0,2 | 0,4 | 6 | fast | slow | Efficiency |
| 10 | 2,0 | 1,6 | 1,3 | 6 | inventive | conventional | Novelty |
| 11 | 0,8 | 5,8 | 2,4 | 6 | obstructive | supportive | Dependability |
| 12 | 2,0 | 2,4 | 1,5 | 6 | good | bad | Attractiveness |
| 13 | 1,8 | 1,4 | 1,2 | 6 | complicated | easy | Perspicuity |
| 14 | 1,3 | 5,1 | 2,3 | 6 | unlikable | pleasing | Attractiveness |
| 15 | 2,0 | 1,6 | 1,3 | 6 | usual | leading edge | Novelty |
| 16 | 1,2 | 4,6 | 2,1 | 6 | unpleasant | pleasant | Attractiveness |
| 17 | 1,2 | 5,4 | 2,3 | 6 | secure | not secure | Dependability |
| 18 | 1,2 | 1,4 | 1,2 | 6 | motivating | demotivating | Stimulation |
| 19 | 1,2 | 2,2 | 1,5 | 6 | meets expectations | does not meet expectations | Dependability |
| 20 | 1,8 | 0,2 | 0,4 | 6 | inefficient | efficient | Efficiency |
| 21 | 2,0 | 1,6 | 1,3 | 6 | clear | confusing | Perspicuity |
| 22 | 1,8 | 1,0 | 1,0 | 6 | impractical | practical | Efficiency |
| 23 | 2,7 | 0,7 | 0,8 | 6 | organized | cluttered | Efficiency |
| 24 | 1,5 | 1,1 | 1,0 | 6 | attractive | unattractive | Attractiveness |
| 25 | 1,7 | 1,1 | 1,0 | 6 | friendly | unfriendly | Attractiveness |
| 26 | 2,0 | 0,4 | 0,6 | 6 | conservative | innovative | Novelty |




| UEQ Scales (Mean and Variance) | | |
|---|---|---|
| **Attractiveness** | 1,528 | 1,47 |
| **Perspicuity** | 1,625 | 1,07 |
| **Efficiency** | 2,292 | 0,21 |
| **Dependability** | 1,000 | 1,73 |
| **Stimulation** | 1,667 | 0,92 |
| **Novelty** | 1,833 | 0,39 |




| Pragmatic and Hedonic Quality | |
|---|---|
| Attractiveness | 1,53 |
| Pragmatic Quality | 1,64 |
| Hedonic Quality | 1,75 |

The scales of the UEQ can be grouped into pragmatic quality (Perspicuity, Efficiency, Dependability) and hedonic quality (Stimulation, Originality). Pragmatic quality describes task related quality aspects, hedonic quality the non-task related quality aspects. Below the mean of the three pragmatic and hedonic quality aspects is calculated.