

**“Wanna talk about it?”:  
Designing conversational AI interactions to  
support lonely young adults**

Masters Thesis

Lauren Beehler  
l.a.beehler@students.uu.nl

July 2024

1st supervisor: Hanna Hauptmann  
2nd supervisor: Christof van Nimwegen

Human Computer Interaction  
Department of Information Science  
Faculty of Science



**Universiteit  
Utrecht**

# Contents

<b>Abstract</b>	<b>6</b>
<b>1 Important Terms</b>	<b>7</b>
<b>2 Introduction</b>	<b>9</b>
2.1 Research Questions . . . . .	10
<b>3 Literature Review</b>	<b>12</b>
3.1 Methods . . . . .	12
3.2 Components of Loneliness . . . . .	12
3.2.1 Types of Loneliness . . . . .	13
3.2.2 Causes . . . . .	13
3.2.3 Affected Population . . . . .	15
3.2.4 Measurement of Loneliness . . . . .	15
3.2.5 Effects of Loneliness . . . . .	16
3.3 Models of Loneliness . . . . .	18
3.3.1 Conceptual Model of Loneliness . . . . .	18
3.3.2 Maintenance Model of Chronic Loneliness . . . . .	19
3.4 Managing Loneliness . . . . .	20
3.4.1 Conversational and Written Therapies . . . . .	21
3.4.2 Cognitive Behavioral Therapy . . . . .	22
3.4.3 The Unified Protocol . . . . .	22
3.5 Digital and Internet-Based Therapy . . . . .	23
3.6 Conversational Agents in Mental Healthcare . . . . .	24
3.6.1 Large Language Models and Conversational Agents . . . . .	24
3.6.2 Application . . . . .	25
3.6.3 Risks . . . . .	26
3.7 Digital Therapies for Loneliness . . . . .	27
3.8 Gap in the Literature . . . . .	28
<b>4 Study Design</b>	<b>30</b>
4.1 User-Centered Design Methodology . . . . .	30
4.2 Participants . . . . .	30
4.2.1 Recruitment . . . . .	32
4.3 Materials . . . . .	33
<b>5 Study 1: Interviews</b>	<b>34</b>
5.1 Materials . . . . .	34

5.2	Data collection, preparation and analysis . . . . .	35
5.3	Results . . . . .	37
5.3.1	Interview Screening . . . . .	37
5.3.2	Interview Participants . . . . .	37
5.3.3	Psychometric Data . . . . .	38
5.3.4	Common codes and requirements . . . . .	39
5.4	Analysis . . . . .	45
5.4.1	Codes . . . . .	45
5.4.2	Profiles . . . . .	50
5.4.3	Profile-based Code Analysis . . . . .	54
5.4.4	Connecting profiles and CBT strategies . . . . .	61
<b>6</b>	<b>Development for Testing</b>	<b>64</b>
6.1	Interaction . . . . .	64
6.1.1	Materials . . . . .	65
6.2	Profile Statements . . . . .	66
6.3	Persona Creation . . . . .	67
6.4	System Prompts . . . . .	67
6.4.1	System Prompt Demonstration . . . . .	69
<b>7</b>	<b>Study 2: User Testing</b>	<b>71</b>
7.1	Materials . . . . .	71
7.2	Data collection . . . . .	72
7.3	Data preparation & Analysis plan . . . . .	73
7.4	Results . . . . .	74
7.4.1	User Testing Screening . . . . .	74
7.4.2	User Testing Participants . . . . .	74
7.4.3	Psychometric Data . . . . .	75
7.4.4	Profile-based statements . . . . .	76
7.4.5	Common codes and feedback . . . . .	77
7.4.6	Interaction selection . . . . .	81
7.4.7	Interaction use . . . . .	83
7.4.8	Interaction evaluation . . . . .	86
7.5	Analysis . . . . .	89
7.5.1	Profiles . . . . .	89
7.5.2	Profile Statements . . . . .	92
7.5.3	Interaction selection . . . . .	92
7.5.4	Experience . . . . .	94
7.5.5	Evaluation . . . . .	95
<b>8</b>	<b>Discussion</b>	<b>96</b>

8.1	RQ1 . . . . .	96
8.2	RQ2 . . . . .	100
8.3	RQ3 . . . . .	102
8.4	Research Contribution . . . . .	105
8.5	Limitations . . . . .	106
8.6	Future Research . . . . .	108
<b>9</b>	<b>Conclusion</b>	<b>109</b>
<b>10</b>	<b>Acknowledgements</b>	<b>110</b>
	<b>References</b>	<b>111</b>
<b>A</b>	<b>Recruitment Poster</b>	<b>126</b>
<b>B</b>	<b>Screening Survey Questions</b>	<b>128</b>
<b>C</b>	<b>Rejection &amp; Mental Health Resources Email</b>	<b>133</b>
<b>D</b>	<b>Information Sheet and Consent Form</b>	<b>136</b>
<b>E</b>	<b>Psychometric Questionnaire</b>	<b>139</b>
<b>F</b>	<b>Interview Script and Questions</b>	<b>144</b>
<b>G</b>	<b>Demographic information</b>	<b>147</b>
<b>H</b>	<b>Profile-based Statement Predictions and Results</b>	<b>148</b>
<b>I</b>	<b>Psychometric Scores</b>	<b>150</b>
<b>J</b>	<b>System Prompt Iteration</b>	<b>151</b>
<b>K</b>	<b>Demonstration Responses</b>	<b>153</b>
K.1	Anne’s Interaction . . . . .	153
K.2	Shou’s Interaction . . . . .	154
K.3	Tanja’s Interaction . . . . .	154
K.4	Jessica’s Interaction . . . . .	155
K.5	Alejandro’s Interaction . . . . .	156
<b>L</b>	<b>User Testing Protocol</b>	<b>157</b>
<b>M</b>	<b>Evaluation Data</b>	<b>162</b>



## Abstract

Loneliness, especially when a chronic issue, can have a large effect on an individual's mental health. The reasons and contributing factors vary, and dimensions of loneliness, including chronicity, permanence, and intensity, can be different for each person and over time. Research has shown that cognitive behavioral therapy has had a moderate effect size for many populations of lonely people. This research seeks to explore how to give people the most effective support by pairing them with established CBT-based interventions based on their experiences and the causes behind their loneliness, based on psychometric connections.

To do so, interviews with nine non-clinical lonely people between the ages of 18 to 40 were conducted to identify needs and perspectives on loneliness. A set of LLM-based conversational interactions was developed that provided interventions depending on the participant's connection to personas based on profiles identified in the interviews. A further five participants tested these CA interactions and their opinions were connected to the profile they were assigned based on their psychometric data.

Results suggest that lonely young adults have high expectations for CA interactions and, when reached, appreciate the ability to discuss their situation and reflect on their thought patterns. Preferred conversation patterns are discussed and connected with strategies to implement them through LLM system prompts. The profiles employed in this research were not comprehensive enough to offer meaningful insights, but qualitative differences inform user expectations and preferences.

# 1 Important Terms

## Acronyms

---

CA	Conversational Agent
CBT	Cognitive Behavioral Therapy
CR	Cognitive Restructuring
GAD-7	Generalized Anxiety Disorder, 7-item
LLM	Large Language Model
PANAS	Positive and Negative Affect Schedule
PHQ-9	Patient Health Questionnaire, 9-item
ppHMS	Patient-Perceived Helpfulness of Measures Scale
SIAS-6	Social Interaction Anxiety Scale, 6-item
SPS-6	Social Phobia Scale, 6-item
UCD	User-Centered Design
UCLA-LS-9	University of California Los Angeles Loneliness Scale, 9-item
UEQ	User Experience Questionnaire
UES-SF	User Engagement Scale, Short Form

## Glossary

---

Term	Definition
(Positive/Negative) Affect	Mood or emotional state; positive and negative affect are two different scales, not two ends of a single spectrum <sup>1</sup> .
Depression	A persistent mental health condition associated with sad, empty, or irritable moods and somatic and cognitive changes that impact an individual's ability to function <sup>2</sup> .
Generalized Anxiety Disorder	A persistent mental health condition associated with excessive fear and anxiety about multiple events and activities, which is difficult to control and impacts an individual's ability to function <sup>2</sup> .
LLaMa 3 8B	An open-source large language model with 8 billion parameters. It was developed by Meta, designed for commercial and research use in English, and released on April 18, 2024 <sup>3</sup> .
Loneliness	A distressing emotional experience caused by a mismatch between a person's perceived and desired social relationships <sup>4</sup> .

<b>Term</b>	<b>Definition</b>
Maintenance Model	An outline of the factors and phenomena that contribute to the continuation of psychopathy or a psychological experience <sup>5</sup> .
Percentile	A value which indicates the percentage of data points in a population that falls below a specific data point; a score $i$ is at the $k$ th percentile if $i$ is greater than $k\%$ of the population's scores <sup>6</sup> .
Persona	A fictitious character modeled on a certain sub-population, often including demographic, personality, and motivational elements <sup>7</sup> .
Profile	A subpopulation with common characteristics (i.e. psychometric ranges) <sup>8</sup> .
Psychological Interventions	Actions or strategies intended to help individuals dealing with psychological conditions, often based on interpretations of maintenance mechanisms and psychological frameworks <sup>9</sup> .
Psychometrics	Tools used to quantitatively measure psychological phenomena <sup>10</sup> .
Social Anxiety Disorder	A persistent mental health condition associated with anxiety about social interaction in dyads or groups <sup>11</sup> .
Social Phobia	A persistent mental health condition associated with fear about social situations where the individual's social performance is exposed to possible scrutiny by others <sup>11</sup> .
System Prompt	An engineered series of statements that describe the intended tones, behaviors, and actions of an LLM <sup>12</sup> .



## 2 Introduction

Most, if not all, people have experienced the sting of loneliness at least once. Loneliness hinges on the mismatch between desired and perceived relationships and social situations. Thus, loneliness can plague people with healthy and supportive social networks and rarely bother people in near isolation. Loneliness is a complex and difficult phenomenon to live with; its experience can impede daily activities and desires, potentially leaving people even lonelier. When left untreated for a long period, chronic loneliness can also have negative effects on psychological, cognitive, and physical health.

Unfortunately, there is a stigma associated with experiencing and expressing loneliness which can make loneliness more extreme and often deters people from seeking help<sup>13</sup>. The stigma, which materializes in the words of others and self-talk, is stronger for young people and seen as less controllable for people living in more individualistic countries<sup>14</sup>. Additionally, lonelier people tend to blame their loneliness on internal reasons, feel more shame, and are more inclined to hide their feelings of loneliness<sup>14</sup>.

In 2023, the U.S. Surgeon General listed Social Connection as one of the top six public health issues and stressed how loneliness and isolation "represent profound threats to our health and well-being,"<sup>15</sup>. The U.K.'s Minister for Loneliness, a position since 2018, has funded research on the topic and focuses on reducing the stigma associated with loneliness<sup>16</sup>. The Netherlands has created groups and campaigns, including the *Nationale Coalitie tegen Eenzaamheid* (National Coalition against Loneliness) and *Een tegen Eenzaamheid* (One against Loneliness), to reach out to and support people dealing with loneliness, supported by a budget of over €4.6 million<sup>17</sup>. This governmental support and focus on loneliness is indicative of the impact of loneliness on people, and an awareness of the dangers of unmanaged loneliness.

However, mental health resources fail to meet the demands. In all countries, the treatment gap, or the proportion of all people who would benefit from treatment compared to those who have access is over 50%, and over 90% in the countries with the least resources<sup>18</sup>. Social and economic contexts, including socioeconomic status, gender, geography, racial and ethnic background, and sexual orientation, can all play a role in determining who has access to mental health care<sup>19</sup>. As the deficiency of access typically affects those who are discriminated against, this lack of access can further stress and limit marginalized people<sup>20</sup>.

Scalable, equitably accessible resources are needed to address this gap; one solution is the use of free digital mental health applications that adjust to user needs. Resources like this enable personalized support and may be as effective as some in-person therapy. According to a recent meta-analysis of digital inter-

ventions for loneliness across all age groups, significant evidence showed that relational agents, or software entities that interact and build relationships with users, could moderately reduce loneliness<sup>21</sup>. However, the relevant studies that included participants under 50 years old were limited, and none of them exclusively focused on this younger population.

The use of conversational agents, while posed to counter the engagement and scalability issues of existing, non-adaptive interventions, has barely been studied within younger adults. Based on the differences in perception and use of conversational agents between younger adults and older adults, one study recommended designers of conversational agents adjust the features based on the age of the user<sup>22</sup>. Studying the differences in desires for, reactions to, and effects of conversational agent interventions of younger adults can improve these resources<sup>23</sup>. The research related to digital interventions tends to be focused on certain diagnoses (e.g., depression, general anxiety) or special user groups (e.g., carers, people with chronic illnesses). Digital conversational therapies that directly focus on loneliness and related topics tend to address older populations.

This paper will build on this dearth of research, addressing the loneliness of younger adults (from 18 to 40) and how conversational agents can be tailored to this population. The desires and characteristics of this population will be explored, conversational interventions will be developed, and the quality of the interventions will be evaluated. Increasing research on this important concern for a population often disregarded with emerging technology is a motivational combination, and this research can inform development, innovation, and future study of this valuable intersection.

## 2.1 Research Questions

In this thesis, I develop LLM-based conversational interactions based on CBT-based interventions for loneliness. This research utilizes primarily qualitative research methods but also explores the use of psychometrics to personalize CA interactions. My research addressed the following Research Questions (RQs) and sub-questions:

1. How do lonely young adults want to interact with a conversational agent (CA) when discussing loneliness?
  - (a) What are their expectations for and desires regarding the interaction?
  - (b) What conversational elements do they want in conversations about loneliness with CAs?
  - (c) What concerns do they have about discussing their mental health with CAs?

- (d) How much does psychometric similarity impact loneliness experiences?
  - (e) What user characteristics can be used to discern meaningful differences in the desired CA experience?
2. How can a text-based CA based on a large language model (LLM) support lonely young adults?
    - (a) What interaction components are most important?
    - (b) What elements should be included in an LLM system prompt for this population?
  3. How are CBT-based CA interventions informed by users' psychometrics perceived by lonely young adults?
    - (a) How effectively can CA system prompts apply the intended loneliness interventions?
    - (b) To what extent do different system prompts offer participants a different experience?
    - (c) Does the use of profiles offer adequate personalization to users?
    - (d) How do users perceive the CA and their interaction experiences?

RQ1 and RQ2 are primarily exploratory RQs, focused on identifying, grouping, and evaluating relevant factors and experiences. RQ1 is answered through interviews with the target population and confirmed during user testing. RQ2 is informed by the interviews and realized through prototyping, testing, iterating, and evaluating a CA interaction. RQ3 is primarily evaluative, tackled during user testing, and includes a focus on the success of the interaction and interventions, judged in part by evaluations of usability and acceptance.

## 3 Literature Review

The fields of loneliness, mental health interventions, conversational agents, and their intersection are full of recent, valuable research. The following literature review aims to summarize the most important elements of these areas.

### 3.1 Methods

The literature review required to address the research questions was conducted first by seeking the most prominent research related to *loneliness*, its causes, the demographics it affects, its comorbidity with other psychiatric disorders, and psychometrics used to measure loneliness. Additionally, I sought leading papers related to *interventions for loneliness*, as well as interventions for depression, general anxiety, and social anxiety, due to the large overlap in the population and treatment. I explored therapeutic research, especially that which was completed using computers, primarily *internet-based CBT*, *digital journaling*, and other implementations of the interventions I had found. To address the interaction I will make, I also searched for important papers related to *large language models* and *conversational agents*, and qualities of the *dialog interaction* that affect user experience. From there, backward snowballing was used on the most relevant papers, and forward snowballing was used to identify the most recent research.

### 3.2 Components of Loneliness

Humans, like many other animals, are social creatures that rely on others to survive; this biological reality is reflected in how many people experience a lack of social relationships. Loneliness is a feeling most people feel at different points in their lives. In contrast to the objective phenomenon of being alone, loneliness is the subjective, distressing experience of not having the desired quantity or, primarily, the desired quality of relationships<sup>4</sup>. While much research exists on individual's definitions of desired relationship qualities, their perception of their relationships, and the factors that affect both of those concepts, this is missing a key element required for the analysis of loneliness. The cognitive processes related to the development of these personal concepts and attributing meaning to the gap between them are essential in the experience of loneliness<sup>13</sup>. Without maladaptive cogitations, a gap between perception and desire could be noticed and acted on. However, when negative interpersonal appraisals, emotional responses, and self-focused cognitive processes are added to the mix, the gap gains blame, shame, sadness, and other distressing emotions<sup>5</sup>. Loneliness is correlated and shares similar cognitive processes with depression, general anxiety, social anxi-

ety, and other somatic disorders, but it is a distinct and independent experience<sup>24</sup>.

### **3.2.1 Types of Loneliness**

There are many different ways to categorize types of loneliness. One of the earlier research-backed dimensions of loneliness, developed by de Jong-Gierveld in 1973, distinguished the emotional characteristics, the "type of deprivation", understood as the quality or type of social relationship a participant desires but does not have, and the time perspective, or how permanent an individual felt their loneliness was<sup>13</sup>. Her study identified four profiles found in the population, primarily based on an individual's loneliness and time perspective: the nonlonely group, the dissatisfied and hopeless lonely group, the periodically and temporarily lonely group, and the resigned lonely group<sup>13</sup>.

The research by de Jong-Gierveld, however, only focused on emotional loneliness, which is the feeling of missing deep, emotional connections and attachments within existing relationships, including feelings of intimacy and feeling 'known' by others<sup>25</sup>. This is contrasted with social loneliness, which instead arises due to a perceived lack of social network, including people to spend time with, support, or get help from<sup>25</sup>. Weiss<sup>26</sup> describes emotional loneliness as a longing for deep attachment and social loneliness as feeling left out or excluded.

Further distinctions about loneliness relate to the amount of time that someone experiences loneliness. Transient loneliness is the most common type of loneliness and is felt by most people. Transient loneliness is a short-term experience, typically a few days or less, while situational loneliness is often caused by larger life events and lasts on the order of weeks or months<sup>13</sup>. Transiently and situationally lonely people learn to accept and deal with their situation or loss and recover from loneliness with time<sup>13</sup>. Chronic loneliness is closer to a trait or characteristic of someone's personality, and typically requires at least two years of loneliness to be categorized as such<sup>13</sup>. Situational loneliness, when not addressed, can transition into chronic loneliness following the habitual use of maladaptive cognitions. The differences between these categories of loneliness are primarily related to the chronicity of experience, the most effective interventions, and the pervasiveness of loneliness in one's life, but not the strength of the emotion<sup>13</sup>. However, the effects of these categories are distinct: chronic loneliness is the cause of the physiological, cognitive, and mental health effects described in Section 3.2.5.

### **3.2.2 Causes**

Loneliness is known as a transdiagnostic experience, present across diagnoses and causes<sup>27</sup>. These causes can be broadly delineated based on someone's relational, intrapersonal, situational, and cultural characteristics<sup>28</sup>, but loneliness

is more likely to be caused by combinations of elements from each category. Each cause and influencing factor can impart different characteristics to the experience of loneliness. For example, loneliness due to the loss of someone close leads to different behaviors, emotions, and cognitions around loneliness than someone whose loneliness revolves around how their social group does not match their values and they feel that no one "gets" them.

Although solitude can be enjoyable if the perceived and desired social relationships align, one cause of loneliness is the objective lack or loss of valued social relationships. Heu et al.<sup>28</sup> found many relational or interpersonal causes stemmed from specific relationship domains or were directly connected to certain contexts. This could include the lack of a certain relationship type (e.g., romantic relationship, specific familial role, best friend), a separation from valued individuals (due to distance, logistical disconnect, or death), and a missing quality within existing relationships, like true understanding, commitment, protection, or vulnerability. Other common and cross-cultural situational causes of loneliness include significant events like moving away from existing social networks, going through a large life shift like losing a job or starting a new job, having a child, gaining a disability, or losing access to connecting technologies (e.g., transit, communication devices)<sup>28</sup>. Whether or not someone has a romantic partner and their overall health are the two most important predictors of loneliness in all age groups<sup>29</sup>.

Genetics have been found to account for approximately half of individual differences in perceived social isolation, presumably reflecting variations in sensitivity to the pain of social disconnection, not the intensity of the pain itself<sup>30</sup>. Loneliness is also influenced by physical and mental disabilities or differences, childhood experiences, cultural norms, social norms for your life stage, gender, and income<sup>4,28,31</sup>.

Finally, research has identified rising levels of loneliness over the last decades, primarily among young people<sup>8</sup>. There are many possible reasons for this rise, including the different environmental factors that could affect loneliness levels. According to a systematic literature review from O'Day et al.<sup>32</sup>, people with social anxiety or loneliness have a higher risk of passive, frequent social media use, and lonely individuals are more likely to experience problematic social media use, especially when engaged in social comparison. Although problematic social media use did not predict loneliness, lonely adults' problematic social media use can be driven by maladaptive cognitions, a higher need for social assurance, and a desire to "escape" feelings of loneliness<sup>32</sup>. Additionally, the COVID-19 pandemic and the involved lockdowns were found to have led to a small average increase in loneliness<sup>33</sup>. Other factors, including the physical and social environment, political upheaval, and changes to other societal norms and expectations, could also affect perceptions of loneliness<sup>34</sup>.

### **3.2.3 Affected Population**

Anyone can be lonely, and most people have been lonely for at least short periods of their lives. Weiss<sup>26</sup>, one of the first prominent researchers of loneliness, stated in his 1973 book, "There is no time in one's life when loneliness ceases to be a threat."

When looking at populations in general, loneliness levels vary. Among young adults in northern Europe from 2000 to 2019, estimates of chronic loneliness ranged from 1.8% to 4.5% and tended to be higher in eastern Europe<sup>35</sup>. However, in 2022, the Netherlands' Central Bureau of Statistics found that 12.3% of people between 15 and 45 experienced strong loneliness, with another 29.8% of that age range falling within the somewhat lonely classification<sup>36</sup>.

Since loneliness is the difference between someone's perceived and desired relationships, an individual's definition of their desired relationships is an important factor that can affect their experience of loneliness. There are differences in the norms that different groups abide by that play a role in how socially engaged an individual believes they "should" be. One of the most powerful sets of norms for social interaction is determined by one's age. Young adults may be instilled with the sense that this is the time in their lives when they should be making friends, setting up their careers, and starting a family, so a single young adult with no job and few friends may experience more loneliness than an older adult with the same characteristics<sup>31</sup>.

In a meta-review of 582 studies, Maes et al.<sup>37</sup> found a significant difference between men's and women's loneliness levels when in young adulthood; with a small effect size, men of this age are lonelier than their peers. Some research has indicated that whether or not the psychometric measure of loneliness is direct (e.g., asks something similar to "Do you feel lonely?") or indirect (e.g., focuses on aspects of loneliness without using the terms "lonely" or "loneliness") affects loneliness attribution. Nicolaisen and Thorsen<sup>29</sup> found that indirect measures (namely the De Jong Gierveld Scale) indicate more loneliness in men under 50, direct measures indicate more loneliness in women than men, and, with measures of social loneliness, men are lonelier than women in all age groups.

### **3.2.4 Measurement of Loneliness**

When it comes to measuring loneliness, several ways have been suggested but a select few have become popular across research. As mentioned previously, there are both single-item, direct measures of loneliness and multi-item indirect questionnaires, and some demographic elements affect responses gathered by these. The differences are often attributed to a narrow view of loneliness and the possibility of a social desirability bias due to the stigma associated with lone-

liness, but when surveying adults with limited time, single-item measures can still reliably capture loneliness<sup>38</sup>.

However, as loneliness is the focus of this thesis and will be a main exclusion criterion, a multi-item measure is preferred. The most common multi-item psychometric measure for loneliness is the University of California Los Angeles Loneliness Scale (UCLA-LS), used by nearly 65% of the 2,318 reports in the Meta-Analytic Study of Loneliness database<sup>25</sup>. The original UCLA-LS uses 20 items and four response categories from 1 (never) to 4 (often), and focuses on the frequency of loneliness experiences; multiple versions of the UCLA-LS are available, each shortened to a total of three, four, eight, or nine items<sup>25</sup>.

Originally designed as a three-factor model, the nine-item UCLA-LS's factors have not held up to scrutiny; however, this version achieved full scalar invariance as a full measure and had the best measurement invariance of all the UCLA-LS version across age groups<sup>39</sup>. An additional benefit of the UCLA-LS-9 is the ability to also compare it to the UCLA-LS-3, which makes up three of the nine items. For UCLA-LS-9 between the ages of 18 to 45, the mean score is 18.31 and the standard deviation is 6.08<sup>39</sup>.

### **3.2.5 Effects of Loneliness**

Loneliness is a risk factor for multiple types of mortality and is comparable in potency to cigarettes, high blood pressure, obesity, or a sedentary lifestyle<sup>40</sup>. Chronic loneliness accelerates aging over time by lowering physiological resilience, leading to larger differences in cardiovascular and neuroendocrine health for older lonely individuals than their non-lonely peers<sup>41</sup>. Loneliness fluctuations were correlated to the level of cortisol, a hormone indicative of stress levels<sup>42</sup>. Additionally, lonely people both get less salubrious sleep and have more daytime dysfunction, but being tired can also increase feelings of loneliness<sup>42</sup>. And if loneliness does not kill or disable you, it can still negatively impact your life beyond the emotional distress.

Loneliness levels can be used to predict cognitive decline and the onset of Alzheimer's disease<sup>43</sup> and lifetime change in IQ<sup>30</sup>. Chronic loneliness in young adults has been found to reduce attention regulation<sup>44</sup> and diminish executive control<sup>4</sup>. Southward et al.<sup>45</sup> found that more severe loneliness was associated with a lower frequency of emotional recognition and emotions' cognitive effects, potentially due to lower motivation or ability. Other research has found that lonely people tend to forecast their future social exclusion and have limited memories of past inclusion<sup>46</sup>.

Gow et al.<sup>47</sup> connected higher levels of loneliness to lower satisfaction with life. Hawkey et al.<sup>41</sup> found that lonely students experienced more stress, helplessness, and feelings of threat than their peers did for the same experiences. Research



has suggested that the methods chronically lonely people use to deal with stress, primarily behavioral withdrawal, tend to perpetuate stress instead of managing the emotional response<sup>41,44</sup>.

One of the most common psychological phenomena of loneliness, however, is the comorbidity of chronic loneliness and depression, general anxiety, and/or social anxiety symptoms. In a study in the Netherlands shortly before 2010, van Beljouw et al.<sup>48</sup> found that in patients with both depression and anxiety, loneliness is the primary factor that predicts poor outcomes after a year, even when the baseline symptom severity is controlled. Other research with older adults from Ireland suggests that loneliness, especially emotional loneliness, could be predicted by depressive symptoms<sup>49</sup>. Further, in a six-month study on adults from the USA, Lim et al.<sup>50</sup> found that loneliness measures could be used to predict levels of social anxiety, paranoia, and depression symptoms, and social anxiety levels could predict future loneliness. Although the relationships between loneliness, depression, and various forms of anxiety are contested, the cooccurrence is evident and can lead to greater effects than any syndrome by itself.

Individuals dealing with chronic loneliness are less likely to engage in regular physical exercise<sup>51</sup> and prioritize healthy foods<sup>52</sup>. As emotional regulation is a key component of self-control behaviors, the diminished ability to emotionally regulate is theorized to lower discipline<sup>42</sup>. The lowered levels of physical activity, typically seen as a protective factor, may increase the chance of poor physical, mental, and cognitive functioning<sup>53</sup>, though animal research suggests that loneliness itself dampens the effects of exercise<sup>42</sup>.

Loneliness has been found to increase emotional reactivity to negative social stimuli and activate the reward center of the brain less with positive social stimuli<sup>4</sup>. This attentional bias and emotive reaction can trigger confirmatory and memory-based biases that can develop into a hypersensitivity to negative social outcomes that motivate lonely people to avoid social situations entirely. They may remember more negative social information and see the social world as a more threatening place<sup>42</sup>.

Humans can perceive social relationships where no objective, reciprocal relationship can exist,<sup>42</sup>. Chronically lonely people may manufacture replacements for social relationships, including reliance on nostalgia, parasocial relationships, and anthropomorphism of pets, technological devices, and supernatural entities<sup>4</sup>.

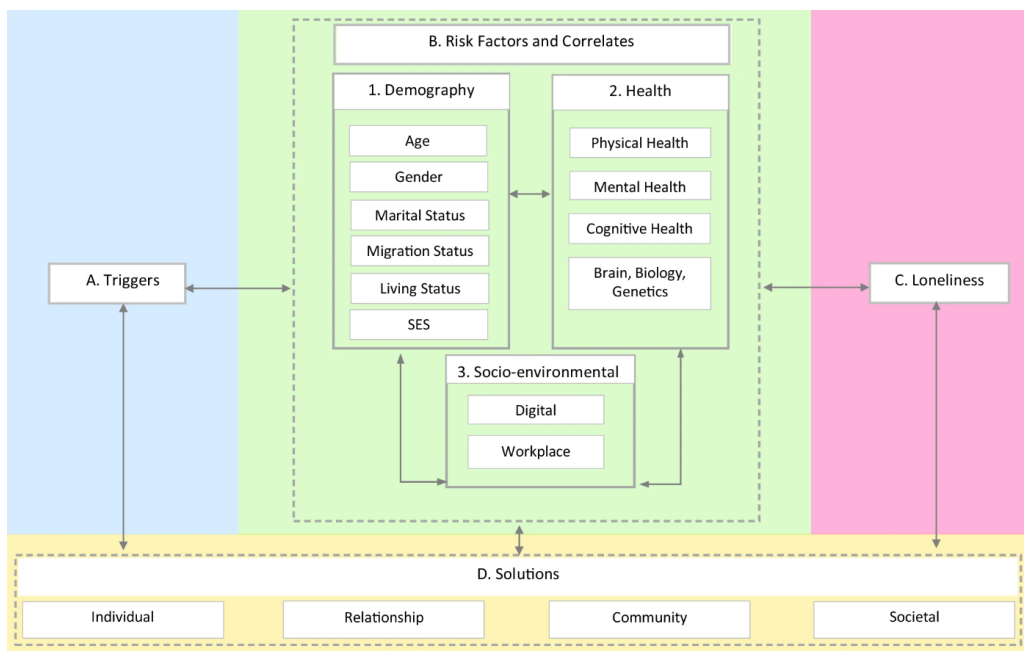
All of these effects of loneliness feed into each other and can lead to self-defeating interactions. Feelings of loneliness increase awareness of the self and sensitivity to social rejection, lower self-esteem, physical health, and mental health, and increase the desire to avoid social situations, which perpetuates the cycle and increases feelings of loneliness. For chronically lonely individuals, loneliness can feel like a permanent state and be very difficult to escape from without help.

### 3.3 Models of Loneliness

To understand the diversity of causes, experiences, and mechanisms that prolong loneliness, several models have been developed. Within this thesis' focus, two models are helpful and will be utilized when developing the resultant interaction. Lim's conceptual model of loneliness outlines the contextual elements surrounding loneliness and Käll's maintenance model of chronic loneliness features the elements that promote the experience of loneliness within an individual.

#### 3.3.1 Conceptual Model of Loneliness

Lim et al.<sup>34</sup> developed a conceptual model of loneliness meant to highlight the complexity of the causes of loneliness, shown in Figure 1. Lim developed this model, which focuses on A) triggers, B) loneliness risk factors, C) the experience of loneliness, and D) solutions, based on literature since 2006 about loneliness and its correlated factors.



**Figure 1.** Lim's conceptual model of loneliness<sup>34</sup>. SES = Socio-economic status.

Lim identifies that triggers are significant life events or transitions in life stages that involve a change in social identity; these moments elevate the risk of experiencing problematic loneliness. The risk factors and correlates are an overview of the fixed and modifiable factors that can increase someone's risk of loneliness. However, the relevance of the risk factors varies by individual and

trigger. The loneliness experience includes dimensions related to severity, pervasiveness, chronicity, and perceived permanence, resulting in different experiences between individuals. Finally, Lim uses the CDC's socio-ecological framework of solutions (originally developed for preventing violence)<sup>54</sup>, which groups solutions into individual interventions, relationship support, community building, and societal efforts.

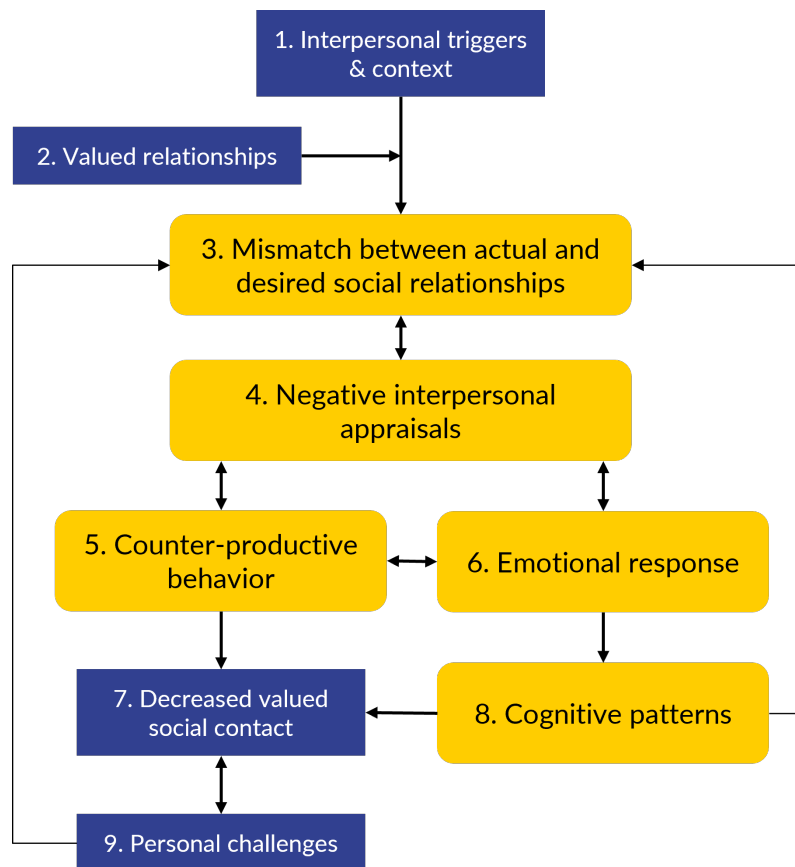
The four-part structure (Triggers, Risk Factors & Correlates, Loneliness Experience, and Solutions) is a helpful way to consider how the context of someone's life can lead to changes in loneliness and identifies some of the components that people can point to as causes and thus address. In this thesis, these elements will inform lines of inquiry with users and likely be brought up to users in the interactions developed.

### **3.3.2 Maintenance Model of Chronic Loneliness**

When analyzing the use of different interventions for the treatment of chronic loneliness in literature, Käll et al.<sup>5</sup> clarified the maintenance factors of chronic loneliness. He hoped this maintenance model, seen in Figure 2, would help people develop theories related to loneliness and enable more efficacious interventions. The model of maintenance for chronic loneliness enables a modular view of interventions based on individual tendencies and challenges. This framework is based on the social anxiety maintenance model, which also has mechanisms related to avoidance of social contact and self-focused attention. Käll points out that interventions based on this model should be introduced in the context of existing social, societal, and specific-group interventions for loneliness as the methods are complementary.

Within the model, some of the key mechanisms include the perceived mismatch between the actual and desired social situation, the negative interpersonal appraisals, counterproductive behavior, emotional response, and cognitive processes. Käll points out that the model shows how transient and situational loneliness can transform into chronic loneliness based on the maladaptive cognitions someone deals with. Those with transient and situational loneliness may also experience a mismatch between the actual and desired social situation and emotional response, but the mechanisms in the bottom half of the diagram enable loneliness to persist for a longer time<sup>55</sup>.

For this thesis, this maintenance model will be utilized to aid in identifying the influences of a participant's chronic loneliness and connecting those to the best interventions for them. Käll identified some key connections in the paper in which he developed the model by linking experimentally successful interventions to the different components. For example, to address an emotional response, people can use relaxation, stress management, problem-solving, and emotional awareness to



**Figure 2.** A maintenance model of chronic loneliness based on research by Käll<sup>5</sup>. The yellow boxes are the elements relevant to the connection of profiles and CBT strategies in Section 5.4.4.

understand, deal with, and control their emotions.

### 3.4 Managing Loneliness

Luckily for people with chronic loneliness, there are ways to manage and lower the emotions and experience of loneliness. Additionally, lowering loneliness levels has been found to predict social anxiety<sup>50</sup> and to statistically significantly reduce depression symptoms<sup>56</sup>, making the alleviation of loneliness a possible way to treat or prevent other mental health diagnoses.

When managing loneliness, research has identified four key intervention categories: 1) improving an individual’s social skills, 2) providing support in social situations, 3) providing or encouraging opportunities for social interaction, and 4) addressing maladaptive social cognition<sup>42,57,58</sup>. Qualitative research from the perspective of lonely people has suggested the benefit of higher independence

from others and simply accepting loneliness as a facet of existence<sup>28</sup>; while possibly beneficial for those who feel lonely despite having fulfilling social relationships, these strategies could be manifestations of avoidant behavior and depression and will not help other populations. In a meta-analysis of randomized controlled trials, meditation/mindfulness, social support, and adapting social cognitions were found to be the statistically significant methods of decreasing the perception of loneliness<sup>9</sup>. The most effective treatment has been to address the unhelpful and distressing thoughts that occur in and related to social situations. Attending to maladaptive cognitions has an average effect size for a decrease of loneliness of Cohen's  $d = 0.60$ <sup>58</sup>. One of the most common tools to deal with these maladaptive cognitions is cognitive restructuring (CR), which encourages people to identify unhelpful thoughts, evaluate how well they fit and describe the situation, and modify their cognitions to be more realistic and measured<sup>59</sup>.

However, when paired with a key trauma or primary diagnosis, the treatment of loneliness can become a secondary focus. In these cases, a higher priority given to addressing the emotional and psychological realities can lead to lowered loneliness levels as a side effect. Additionally, when loneliness is present in part due to low social skills, improving these social skills alone through InterPersonal Therapy (IPT) or other types of situational and relational work can be effective<sup>5</sup>. Because of these additional factors, loneliness treatment does not simply consist of reframing thoughts.

### **3.4.1 Conversational and Written Therapies**

Psychotherapies require a relationship between a patient and a 'healer' who is tasked with clarifying symptoms and problems, exploring emotional experiences, and supporting the patient in improving their outlook and perspectives<sup>60</sup>. Therapies can differ in form, content, desired outcome, and rigor, all of which can have strong impacts on the experience. A significant predictor of providing effective therapy is the presence of therapeutic alliance<sup>61</sup>. This quality of a therapeutic interaction requires agreement between the patient and therapist on the therapeutic goals, the development of a sense of trust and connection, and the assignment of therapeutic interventions to be successful<sup>62</sup>.

Conversational therapies are built on the premise that talking about your feelings and problems can help people manage their emotions and determine how they can act while following their values. Talking about emotional events can enable emotional recovery from those events, lessening and adjusting the affective experience when the event is remembered<sup>63</sup>. The standard way of conducting conversational therapy involves the patient and the therapist sitting in the same room and discussing in private, though group therapy can also be effective as an alternative to a waitlist or as an addition to individual therapy<sup>64</sup>.

Unfortunately, conversations about sensitive topics can be difficult, and interpersonal expression may be an intimidating barrier. In these cases, written emotional expression can help produce the positive change that spoken emotional expression offers<sup>65</sup>. One concern of written reflection is that it is intended to lead to emotional processing and possibly distress; while this enables long-term outcomes, unsupported experiences of strong emotions could worsen mental health<sup>65</sup>.

### **3.4.2 Cognitive Behavioral Therapy**

There are multiple potential ways to help someone deal with and correct maladaptive cognitions related to loneliness, but one of the most commonly used ones is Cognitive Behavioral Therapy (CBT). CBT is an active and collaborative treatment meant to reduce emotional distress and increase cognitive flexibility while encouraging patients to collect their own 'data' supporting the efficacy of adaptive strategies<sup>66</sup>. The core CBT therapies, cognitive restructuring, behavioral activation, exposure, and problem-solving, are focused on working through avoidance, increasing social interaction, and planning positive habits and self-care<sup>66</sup>.

As mentioned above, CR teaches people how to determine and label which of their automatic thoughts are causing their distressing feelings, then helps them analyze those thoughts, generate alternative cognitions in that situation, and judge how an alternate thought pattern could affect their emotions<sup>59</sup>. Through Socratic questioning of automatic thoughts, underlying beliefs can also be unearthed; these beliefs can be related to social rejections, personal shortcomings, self-deprecation, and threats<sup>66</sup>. Belief shift tends to happen slowly but, through techniques like a positive data log, role play, and reconsidering situations from an external perspective<sup>66</sup>. In a 2021 meta-review from Erhardt et al.<sup>59</sup>, only 6% of the free applications that advertised cognitive restructuring offered features to adhere to the intervention correctly.

Behavioral action encourages patients' action and engagement in their own lives, primarily through tracking their mastery and pleasure throughout the day and then scheduling activities expected to lift their mood<sup>66</sup>. Some practitioners prompt patients to discern their values and help translate them into meaningful behaviors<sup>66</sup>. Exposure, whether real, imagined, or simulated, demonstrates that exposure to feared stimuli does not need to lead to an adverse outcome and is used repeatedly to reduce the fear response<sup>66</sup>.

### **3.4.3 The Unified Protocol**

The Unified Protocol (UP) was developed using CBT techniques and designed to cover multiple disorders wherein emotional dysregulation is a key mech-

anism<sup>67</sup>. Concurrent and lifetime comorbidities of emotional disorders are 55% and 76%, respectively, allowing the UP to operate on the understanding that they act like different manifestations of the same, broader syndrome and that these disorders can be tackled together<sup>67</sup>. One of the key covariance factors was explained by a higher negative affect and a low positive affect, found to be more frequently experienced by individuals with anxiety and mood disorders<sup>27</sup>. The UP is differentiated from CBT due to its explicit focus on how emotional experiences are built from the interactions between thoughts, feelings, and behaviors, and the introduction to skills and tools that people can use to make emotions easier to manage<sup>27</sup>. Additionally, the UP is adjusted to the emotional symptoms of the patient instead of a specific diagnostic category<sup>67</sup>. Since loneliness is commonly understood to be a transdiagnostic experience primarily maintained through emotional responses and maladaptive cognition<sup>5</sup>, and is correlated with depression, social anxiety, and general anxiety, the UP has also been applied to loneliness<sup>27,45</sup>. Low UP skill frequency was associated with people feeling lonelier than their average, which led to higher depression and anxiety symptoms, but a greater number of UP skills used was able to predict decreases in loneliness<sup>45</sup>.

### **3.5 Digital and Internet-Based Therapy**

The traditional method of therapeutic care is a face-to-face interaction between a licensed therapist and one or more individuals. However, as mentioned in Section 3.4.1, written therapy is also an effective and helpful method of emotional expression and cognitive adjustment. With the advent and proliferation of the internet, therapy was able to move to online spaces. This could appear as synchronous video chats directly between providers and patients, digital self-help resources, and therapeutic exercises that providers can asynchronously assign and review. Previous user-centered design processes related to digital mental health interventions have identified the fundamental elements of convenience, personalization, anonymity, privacy, safety, and non-patronizing, non-judgemental language to be paramount<sup>68</sup>. One of the main differences between internet-based therapies is the access to and of a mental health practitioner.

Guided or blended use is the use of technological intervention in conjunction with more traditional forms of therapy. Guided therapy can have stronger effects and treatment can have higher adherence than in-person therapy alone, allowing therapist contact to be decreased<sup>8</sup>. Some of the benefits of mixing face-to-face therapy and interactive digital tools include the ability to provide helpful prompts to the patient, to help make specific clinical skills more concrete, to track symptoms more reliably, and to enable learning outside of the sessions<sup>69</sup>.

Unguided use is like reading a self-help book; no licensed professional di-

agnoses, makes sure that ideas are interpreted correctly, or aids in tough moments. However, unguided use is typically much cheaper, available 24/7 and wherever the user is, and features many of the same lessons and exercises<sup>20</sup>. Further, apps and chatbots can foster a sense of anonymity and confidentiality that inspires trust and allows people with a stronger within-self stigma to access mental health care<sup>20</sup>.

On-demand regimens enable access to mental health practitioners when needed, but the users spend most of their treatment working through materials and exercises on their own<sup>8</sup>. Additionally, many mental health applications found in app stores are designed to be used on their own, primarily as a self-help resource, but can be appropriated for blended care<sup>59</sup>.

## **3.6 Conversational Agents in Mental Healthcare**

The ability to freely use natural language to describe experiences and for that input to be interpreted accurately is a key element of language-based therapies. To enable that kind of interaction, Conversational Agents (CAs) have been developed to offer more flexibility in conversation and support.

### **3.6.1 Large Language Models and Conversational Agents**

Natural language processing (NLP) is a category of computational processing that analyzes and represents natural human language automatically, utilizing linguistic, computational, cognitive, and communication theories<sup>70</sup>. Language models are NLP models that attempt to interpret the meaning of and generate natural human language based on the predicted next word in a sequence<sup>71</sup>. Large Language Models (LLMs) are language models with many parameters from billions of inputs and can analyze, manage, and create longer text entries<sup>71,72</sup>; LLMs fall within the popular understanding of artificial intelligence (AI). Additionally, when trained on massive corpora of text materials including books, news articles, internet entries, and more, LLMs can utilize in-context learning to handle more than text continuation<sup>72</sup> and can handle more diverse tasks than other machine learning programs<sup>71</sup>.

Current LLMs are proficient at generating fluent and precise text, sentiment analysis, text classification, logical reasoning, and question-answering<sup>71</sup>. Basic operations include content splitting and extraction, rewriting, format matching, and ideation<sup>72</sup>. However, LLMs struggle with semantics, abstract reasoning, filtering out illegitimate and false information, being robust to adversarial prompts, managing the biases of training material, and understanding knowledge, concepts, or terms that are not in the training data<sup>71,73</sup>.

One of the main ways that people interact with LLMs is through prompts or written texts that direct LLMs to complete tasks or develop certain outputs<sup>71</sup>.



Natural language prompts enable users to interact with the LLMs similarly to how they interact with a person. However, when it comes to real-world tasks, LLMs can struggle to identify and complete the intermediary steps or integrate multiple directions<sup>72</sup>.

Colloquially known as chatbots, CAs are language models that focus specifically on dialog interaction and are trained to behave as human-like as possible. The CA must manage the processing of the user input, its combination with contextual information, its understanding of user intent, the integration into the dialog, the information retrieval needed to inform the response, and the response generation<sup>74</sup>. When analyzing and developing the text itself, the CA uses syntactic, semantic, and pragmatic analysis and dialog management, which includes goal handling and inference from prior interactions<sup>75</sup>. Beyond the discussed interaction prompts, system prompts for CAs describe intended tones, behaviors, or actions<sup>12</sup>. System prompts can be hidden from the user and applied to every message from the user, enabling a consistent and intentional “character” or strategy from the CA.

One of the key conversational strategies missing from many CAs is the concept of goal awareness, or the CA’s ability to move a conversation toward a certain goal, but this can be remedied if proactivity and goal use are integrated into the design<sup>76</sup>. Another helpful element that can be merged with CAs is a recommender system, which gives people more control over what is recommended to them<sup>77</sup> and could be used to connect the dialog from the user to the loneliness interventions that would work best for them. A CA that can proactively steer conversations and has a way to connect the user input to certain therapeutic strategies can play the role of a therapist or coach.

### **3.6.2 Application**

LLMs and CAs can enable personalized support and adaptive interactions to help users manage their care more easily. This also enables digital therapeutic alliances, where users pick or agree to suggested tasks that could progress them toward their stated therapeutic goal<sup>78</sup>. Beyond these, CAs can be tailored to the user, making the experience more relevant and effective<sup>20</sup>. For people with chronic diseases, CAs can enable easy, explainable, and consistent monitoring of symptoms and maintenance of the condition with a lower mental load<sup>79</sup>. In a review of the features of chatbots in mental health, most were rule-based (not LLM-based), focused on providing therapy, training, and screening, and were commonly designed for use by those with depression or autism<sup>80</sup>.

Created in 1966, ELIZA was the first chatbot, made to imitate a Rogerian therapist; although it was a rudimentary interaction made of simple dependencies and text substitution, ELIZA was commonly anthropomorphized<sup>81</sup>. Since then,

CAs have advanced tremendously and have enabled human-like conversations and far more personalization.

As a replacement for human therapists, CAs are seen as more trustworthy and obedient, offering care on their terms, enabling users to be more open about their experiences<sup>82</sup>. CAs in mental healthcare are very unlikely to cause adverse events with both clinical and subclinical patients, with an incidence rate of under 0.3%<sup>83</sup>. Patients and mental health professionals were satisfied and confident in the use of CAs, and CAs' ability to promote adherence and for use in psychoeducation were strong suits<sup>83</sup>. When compared with writing in a diary, both interventions had significant impacts on negative emotions but CAs encouraged discussing more sadness and improved overall mood more<sup>84</sup>.

Research about the use of LLM-supported CAs tasked with cognitive restructuring did not need to be trained on human counselors to be effective<sup>85</sup>. Users found that CAs were able to support their desires to "re-author" their lives, or find ethical values and agendas in their thoughts and actions<sup>82</sup>.

### 3.6.3 Risks

However, there are many pitfalls and risks associated with using LLMs and CAs for mental healthcare. Mental health applications, especially those that encourage the disclosure and storage of private, personal data and health information, are in danger of compromised data security. Many mental health applications and CAs do not fall under Health Insurance Portability and Accountability Act (HIPAA) or confidentiality laws in the United States of America, which can result in health data being unsecured or even sold<sup>86</sup>. In Europe, the definition of medical devices includes some mental health applications, typically classified as Class I (lowest risk), but does not apply if the application does not do more than store, archive, communicate, or do simple searches of information<sup>87</sup>. However, Europe's General Data Protection Regulation (GDPR) is a comprehensive protection of all personal data and has special protections for health data<sup>88</sup>. However, there are no laws or other protections to prevent companies from using data gathered from users to train their LLMs, which can be a large risk to the security of health data<sup>89</sup>. Researchers in the intersection of LLMs and privacy suggest that developers of CAs should design the systems to limit security risks and disclosure behaviors and promote informed privacy choices<sup>90</sup>.

One limitation of CAs and other unguided mental health applications is that they cannot complete complex diagnoses or assessments, and some experts argue that equipping users with self-diagnosing tools can lead to incorrect diagnoses or leave users in distress without support<sup>20</sup>. Further, determining the boundary between self-help and professional advice is a grey area that could lead to mistaken interpretations<sup>79</sup>. Additionally, as loneliness and other mental health concerns

can lead to increased parasocial connections and personification of technology, researchers have worried that users may form unhealthy attachments or dependencies on CAs<sup>83</sup>.

Liability is another large concern when dealing with people with mental health issues. Who determines if a user has crossed the line and threatens imminent harm to themselves or others, and who contacts external resources? There are currently no laws or regulations about legal responsibility that guide companies in navigating this area<sup>83</sup>. This issue becomes even more complex as users might communicate with sarcasm or exaggeration that might raise alarm bells preemptively and could cause an application to break confidentiality unnecessarily<sup>20</sup>.

CAs have additional concerns due to the human-like conversations and the mechanics of large language models. Users of mental health software must remember that they are interacting with a computer system instead of a person. Losing sight of that fact, or overestimating the abilities of the digital system to reason, empathize, and suggest treatment is known as therapeutic misconception<sup>20</sup>. Also, since machine learning and large language models "learn" from examples and other existing information, the data that is fed into the model shapes the behavior of the model. If the training data includes harmful biases, incorrect information, or other hidden harm, the model adopts that behavior as well<sup>91</sup>.

### 3.7 Digital Therapies for Loneliness

Wysa<sup>92</sup> is a publicly available application that aims to help users "*experiencing low mood, stress or anxiety, or who are interested in improving their emotional resilience,*" and utilizes free-text, CBT-based conversational agents, described as coaches, and offers Emotional Wellbeing professionals in the premium version. Wysa inspired a therapeutic alliance with users, and users tended to personify the conversational agent, perhaps increasing the users' engagement in the application<sup>93</sup>.

Woebot<sup>94</sup> is another commercial application studied for its efficacy in helping people manage stress, burnout, and resilience. Woebot is grounded in CBT techniques, InterPersonal Therapy strategies, and Dialectical Behavior Therapy, and communicates to the user in text-based dialog. The research found that stress reduction was greater for users with clinical levels of general anxiety disorder or depressive symptoms, but that levels of burnout and resilience increases were statistically similar for clinical and subclinical users.

CareCall<sup>95</sup> was designed to call and monitor socially isolated people to "check-in", and used conversational LLMs to develop human-like questions and answers during conversations. CareCall was deployed to address middle-aged and older adults' loneliness in South Korea, and in a study of people who used,

managed, and developed the system, they described emotional improvement when discussing broad topics. However, the CA struggled with personalization because the memory of the system could not accommodate the personal health history and other details that would have been required to adapt to each user. The system's users were also concerned that the interaction would lead to a decrease in contact with caregivers if the system was seen as a substitute for their monitoring and social interaction.

Further research used the CA Tess<sup>96</sup>, designed to teach coping skills and provide support, to specifically address loneliness and social isolation. They found that most users perceived the interaction positively and again personified the chatbot, giving it qualities like being non-judgemental, caring, and open to listening. The researchers identified a strong belief of users that the interactions were helpful and positive but also found themes of disconnect, negative technical experiences, and a preference for human interactions.

### **3.8 Gap in the Literature**

Based on the literature available, there is a gap to be addressed in this thesis. While CAs have been developed to address loneliness and interventions using CBT and the UP have been effective at alleviating loneliness, there is little research pairing these concepts. In the work that does exist in this intersection, frameworks of loneliness and investigations into the user's context, goals, and psychometrics are not used to inform the conversation. Additionally, further investigation into the user needs and desired experience can inform the development of more comprehensive, engaging, and usable systems. Finally, most of the research related to CAs and loneliness has focused on older adults, a distinctly different population with different needs and expectations surrounding digital interactions. The exploration of all of these topics with this younger population offers a rarely-explored perspective on the experience of loneliness.

This thesis will not include an evaluation of CBT, the UP, or other therapeutic interventions, and it will not involve customizing an LLM or the development of a scalable product. Instead, young adults' experiences with loneliness and their expectations and opinions about LLM-based CA interactions. The knowledge gained is used to identify profiles of users and match them with research-backed therapeutic interventions designed to manage and treat the loneliness of young adults. The resultant user-informed CA interactions are further tested.

The outcomes of this research include qualitative research regarding an often-overlooked population, the creation of a CA and related system prompts, the exploration of psychometric-based profiles related to the experience of loneliness, and recommendations for the conversational styles and features needed to

develop of CA interactions. Finally, this research is a foundation for future research to validate, expand upon, and explore this niche area.

## 4 Study Design

This thesis explores the research questions mentioned in Section 2.1. To answer these, the research methods were focused on the population of lonely young adults and first explored through interviews what they would expect from a digital interaction about their loneliness. Two prototypes were used in the interviews to identify tangible things they appreciated and those they disliked. This research method, referenced as Study 1, is explained in Section 5, the results are introduced (Section 5.3), and these results are analyzed (Section 5.4). Based on these results and findings, an LLM-based interaction guided by the findings of the interviews was developed and evaluated on how well it fit the needs of the population through user testing. The development of the interaction and the other materials required for user testing is described in Section 6. Following this, Study 2 is described (Section 7), the results gathered are shared (Section 7.4), and the analysis of these results follows (Section 7.5). The research questions are then addressed in the Discussion (Section 8), followed by the limitations (Section 8.5) and the potential impact of these studies on future research (Section 8.6). Figure 3 features a diagram of the research's structure.

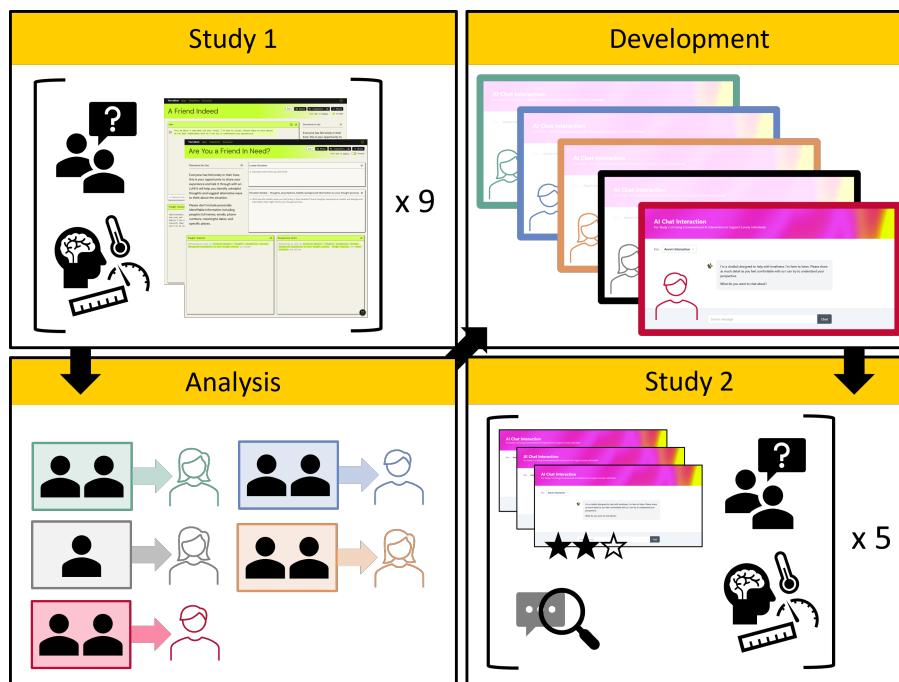
The Ethics and Privacy Quick Scan of the Utrecht University Research Institute of Information and Computing Sciences was conducted and this research was classified as low-risk with no fuller ethics review or privacy assessment required; more details are available in Appendix N.

### 4.1 User-Centered Design Methodology

This thesis utilized user-centered design (UCD), a design process grounded in the extraction and use of information about the relevant stakeholders and users of the final product. The general process of UCD is an iterative loop of learning, developing, and evaluating. The UCD process has been suggested as a good tool for people designing evidence-based treatments in mental healthcare<sup>97</sup>. In this research, the design of the interaction and the user testing to evaluate it was based on the findings of the initial interviews. To address this, the methods and results of Study 1 will be fully explained before introducing the development of the interaction and the methods and results of Study 2.

### 4.2 Participants

The target population is the same throughout the research; I utilized the same inclusion and exclusion characteristics and similar recruitment strategies. This lends itself to better generalizations and the ability to address my research



**Figure 3.** An overview of the research conducted. Study 1, which consists of psychometric evaluation, interviews, and prototype testing, is explained further in Section 5. Analysis describes the identification of psychometric profiles (Section 5.4.2) and the creation of representative personas (Section 6.3). Development of a CA interaction that employs CBT-based interventions matched with the profile is described in Section 6. Finally, Study 2, which uses those targeted interactions as well as psychometric evaluation, interviews, and evaluation, is explained in Section 7.

question for this population. The target population’s characteristics include:

- **Comfortable with English.** The target population speaks, reads, and understands English comfortably, enabling them to understand the research aims, give informed consent, interact with the developed system, and communicate feedback.
- **Aged 18 to 40.** The target population was born between 1984 and 2006, featuring the majority of Millennials and the adult half of Gen Z<sup>98</sup>. This age range was picked due to their high levels of loneliness<sup>31</sup>, a tendency toward better English comprehension skills in the Netherlands<sup>99</sup>, low levels of cognitive impairment outside of depression and anxiety symptomatology<sup>100</sup>, and a higher rate of engagement and comfort with digital mental health interventions<sup>101</sup>.
- **High levels of loneliness.** The participants were selected if they received above-average UCLA-LS-9 scores. The population mean for this age group

was 18.31 in a UCLA measurement invariance report<sup>39</sup>.

- **Non-emergent levels of mental health symptoms.** As this project was not be managed by a psychologist or other mental health professional, people with a present, strong need for psychiatric support were not eligible. This included people who are a danger to themselves (namely those with suicidal thoughts and substance abuse), those who have started or changed their mental health medication in the last three months, and those in active therapy primarily for psychiatric diagnoses.
- **Lives in and around the city of Utrecht, NL.** As the interviews and user testing necessitate in-person interactions, participants must be able to travel to Utrecht or to somewhere within an hour's travel by public transit from Utrecht Centraal.

#### 4.2.1 Recruitment

This research used purposive sampling due to the population description. This research also used convenience sampling due to limitations in time and recruiting resources. To recruit participants from my target population, I used several strategies including in-person communication with friends and acquaintances, social media, and printed posters; one version of the poster is available in Appendix A. The poster was displayed:

- **in subreddits on Reddit:** r/UtrechtExpats, r/UtrechtUniversity
- **in Facebook Groups:** Utrecht Internationals, Utrecht 2024 - International & Erasmus Students, Expats Utrecht, Utrecht International Students, Expats in Utrecht, International Students Utrecht, UTRECHT!
- **in WhatsApp group chats:** Applied Cognitive Psychology, Utrecht Party Group, Marktplaats Utrecht 2024, Mentors & Peer Leaders Welcome Week
- **in Utrecht University buildings:** Universiteitsbibliotheek Uithof, Utrecht Inc, Buys Ballotgebouw, Victor J. Koningsbergergebouw, Educatorium, Ruppertgebouw, Drift 21 & 25, Janskerkhof 3

After sharing information about the research with potential participants in person, online, or via poster, they completed a short questionnaire that helped me verify that they fell within the target population. This survey included the UCLA-LS-9 items, questions about age, gender, and English fluency, and questions exploring if the participant had recently changed psychiatric medication, was in treatment for psychiatric diagnoses, or experienced suicidal thoughts. The



full questionnaire is available in Appendix B. Respondents were then requested to enter their email, enabling further contact.

If respondents were not in the target population, I thanked them for reaching out, commended their proactivity, and offered local and online resources for loneliness, mental health struggles, and suicide prevention; that email is available in Appendix C. The respondents that matched the target population were contacted to schedule interviews and one follow-up email was sent in the case of non-response.

### 4.3 Materials

Some materials were used for both studies and for the organization of the research in general.

- Print and digital versions of recruitment posters (Appendix A)
- Qualtrics XM Survey software (cite Qualtrics XM)
- Screening questionnaire hosted on Qualtrics (Appendix B)
  - England’s Community Life Survey single-item direct measure of loneliness, “How often do you feel lonely?”<sup>102</sup>.  
*The UK’s Office for National Statistics recommends asking this to complement a multi-item, indirect measure of loneliness and enable broader comparisons.*
  - UCLA-LS-9 (explained in Section 3.2.4)
- Psychometric questionnaire hosted on Qualtrics (Appendix E)
  - Depression (Patient Health Questionnaire; PHQ-9)<sup>103</sup>
  - Generalized Anxiety Disorder (GAD-7)<sup>104</sup>
  - Social Interaction Anxiety Scale (SIAS-6), and Social Phobia Scale (SPS-6)<sup>105</sup>
  - Positive and Negative Affect Schedule (PANAS) questionnaire<sup>106</sup>
- Information sheet, consent forms, and resources for mental health support (Copy for interviews is in Appendix D)
- Recording applications (on phone and laptop) and Word Online to create transcripts of the conversations had during interviews and user testing
- NVivo, for open coding and qualitative analysis
- Snacks

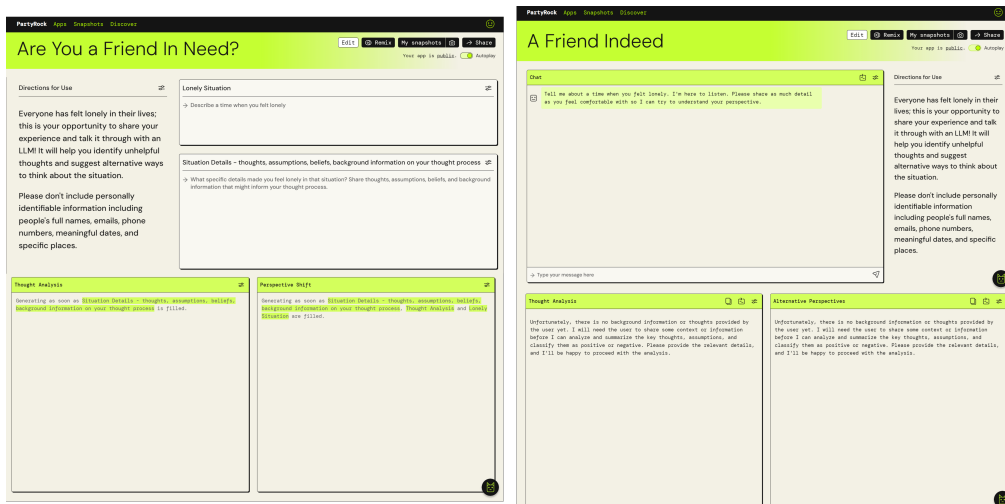
## 5 Study 1: Interviews

The interviews enabled the exploration of the needs and wants of users, the identification of current strategies used to deal with loneliness, the elaboration of preferred interaction methods, a discussion on preconceptions about generative AI in mental healthcare, and more. The interviews were semi-structured to enable flexibility with questioning but ensure that similar topics are targeted across participants. Additionally, to bring the conversation into a more concrete frame and uncover unknown unknowns<sup>107</sup>, the participants had a chance to use prototypes that imitate a rough version of cognitive restructuring.

### 5.1 Materials

In addition to the materials listed in Section 4.3, the following materials were developed for and utilized in the interviews:

- Semi-structured interview protocol (Appendix F).
- PartyRock prototype applications: Are You A Friend in Need? (form interaction) and A Friend Indeed (chat interaction), shown in Figure 4.



**Figure 4.** The empty prototype interfaces participants used; the form interaction is on the left and the chat interaction is on the right.

## 5.2 Data collection, preparation and analysis

Before the interview, each participant filled out a Qualtrics survey containing psychometric items related to depression, generalized anxiety, social anxiety and phobia, and positive and negative affect. The survey took approximately 5 minutes to complete for most participants. The interviews were held in comfortable, relatively quiet spaces and the conversations' audio was recorded with two devices. The text conversations participants had with the prototypes were saved as PDFs. Figure 5 shows an example interaction for each prototype and how those conversations were saved for later reference.

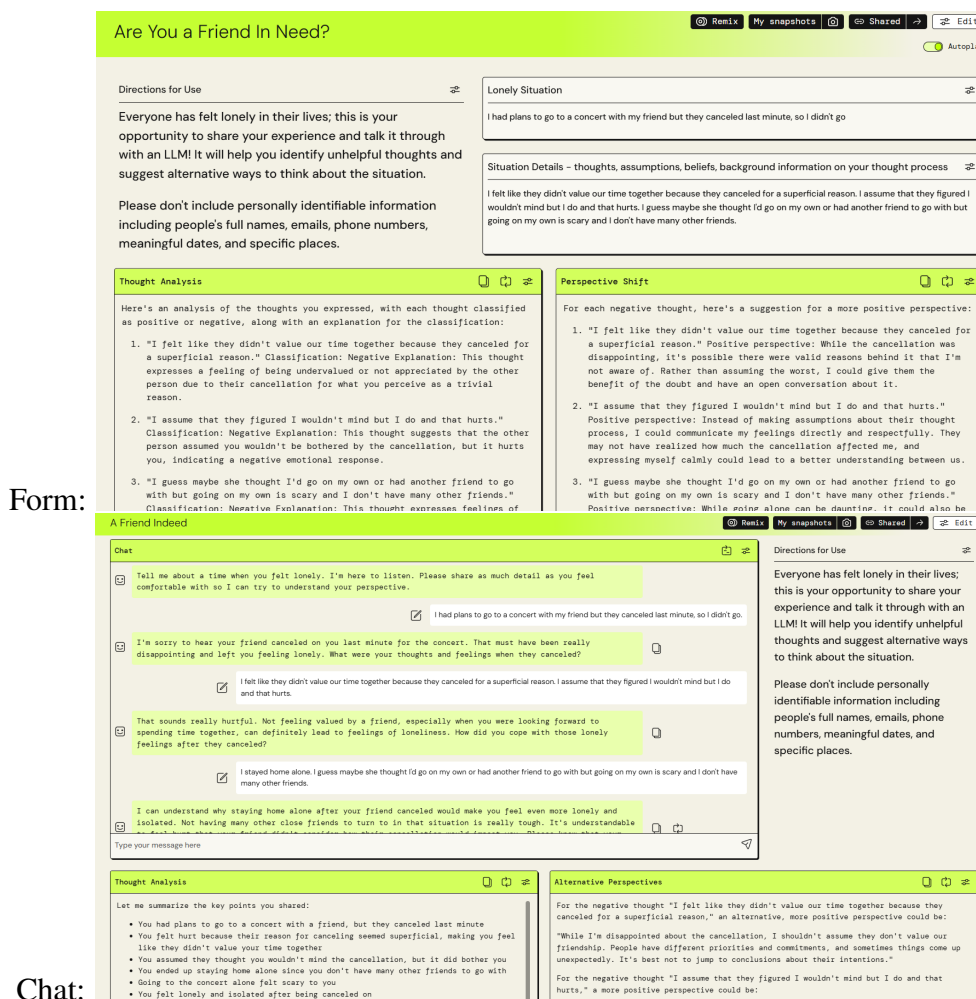


Figure 5. A fictitious conversation with each of the two prototypes.

The transcripts, interactions, and psychometric survey data were pseudonymized

with participant ID numbers. The participants' responses to the psychometric items were transformed into scores according to the guidelines provided with the tools. The mean and standard deviation for the target age group of each psychometric tool were found in confirmatory and population analysis research; these values and their sources are summarized in Table 1. The audio recordings were transcribed using the Word Online transcription tool and corrected manually before deleting identifiable audio. The conversations between the participants and the prototypes were transferred to a Word document.

**Table 1.** Utilized psychometric tests' means, standard deviations, and the sources of these values. The UCLA-LS-9 means, provided in total and for each factor, and the statistically significant difference between GAD-7 values for men and women is reflected in the table.

<b>Psychometric</b>	<b>Mean</b>	<b>St. Dev.</b>	<b>Publication</b>
UCLA-LS-9 F1 (Intimate)	7.74	2.17	Panayiotou et al., 2023 <sup>39</sup>
UCLA-LS-9 F2 (Relational)	4.54	1.88	
UCLA-LS-9 F3 (Collective)	6.03	2.03	
UCLA-LS-9 Total	18.31	6.08	
PHQ-9	15.5	6.0	Sedgwick, 2012 <sup>108</sup>
GAD-7 (women)	7.80	5.80	Byrd-Bredbenner et al., 2021 <sup>109</sup>
GAD-7 (men)	5.44	5.02	
SIAS-6	12.52	5.14	Le Blanc et al., 2014 <sup>11</sup>
SPS-6	10.18	6.01	
PANAS: Positive Affect	33.3	7.2	Crawford and Henry, 2004 <sup>106</sup>
PANAS: Negative Affect	17.4	6.2	

The psychometric and demographic data collected in the screening and psychometric surveys were summarized with descriptive statistics and percentiles. Based on the patterns between the psychometric tools, discernable profiles were identified. Open coding in NVivo was performed on the transcripts to identify patterns in the responses. The codes were grouped into more precise thoughts, needs, and desires. The interactions were also coded in NVivo based on the helpful and unhelpful conversation patterns discussed in the interviews. The profiles, along with demographic data from the screening survey and usage behavior from the prototype interactions, were compared with the opinions and preferences stated in the interviews.

## **5.3 Results**

Nine interviews were successfully held with members of the target population. The psychometric scores of the user testing participants are in Appendix I and the anonymized interview transcripts are available on request.

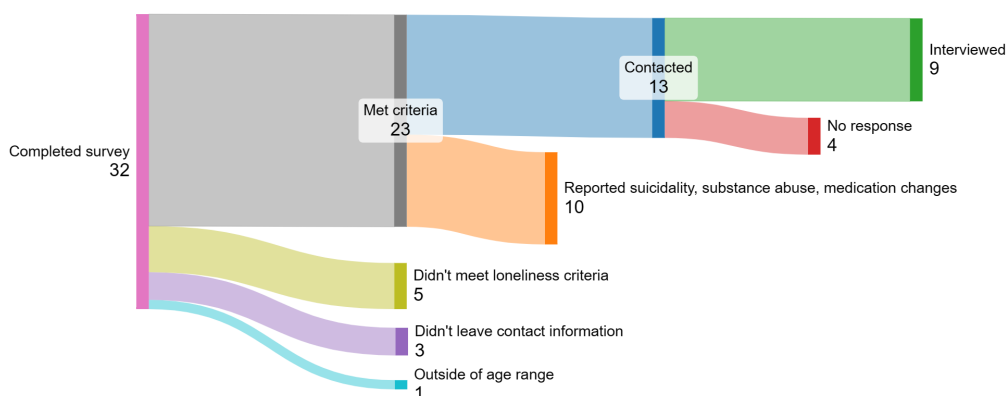
### **5.3.1 Interview Screening**

After sharing the poster containing the link to the recruitment survey online and hanging it in Utrecht, a total of 32 full responses were collected between February 20 and March 26, 2024. This data can illustrate a sample of the help-seeking population. The data can also be compared to the interview participants; the interview participants are not significantly different from the recruitment survey respondents. The UCLA-LS-9 scores ranged from 13 (19th percentile) to 32 (98th percentile), with an average score of 24.31 (SD 5.14). 20 of the respondents identified as women, 11 as men, and 1 did not disclose their gender. They ranged in birth year from 1981 to 2005; the median birth year was 1999 and the mean was approximately 1998. Respondents reported diagnosed or suspected diagnoses of ADHD, autism, BPD, depression, dyspraxia, generalized anxiety disorder, eating disorder, post-traumatic stress disorder, and substance abuse. The respondents' English fluency levels included two B1 speakers, one B2 speaker, six C1 speakers, 16 C2 speakers (making up 50% of respondents), one "I don't know," and six native English speakers.

The plurality of respondents (11) were excluded due to having emergent psychological needs beyond the target population of this research. Nine reported suicidal thoughts or plans, four reported substance abuse issues, and two reported both. Of these respondents, one did not leave their contact information and three others also reported recent changes to their psychiatric medication. Five respondents did not reach the UCLA-LS-9 mean of 18.31, one respondent was a year older than my target population, and three did not leave their contact information that would have been necessary to schedule an interview. The remaining 11 respondents were emailed to set up an interview time; six did not respond to the first email so a follow-up email was sent, but four never responded. An overview of the reasons for excluding respondents is available in Figure 6.

### **5.3.2 Interview Participants**

In March and early April 2024, nine interviews were held with members of the target population. The interviews followed the protocol detailed in Appendix F and featured the prototypes explained in 5.1. Each interview ranged between 55 minutes to 1 hour and 33 minutes. On average, one week elapsed between



*Figure 6. Reasons for exclusion of respondents to the screening survey.*

when the participant completed the screening Qualtrics survey and their interview. Seven interviews were held on Utrecht University’s campus, mostly in classrooms in Buys Ballotgebouw or private rooms in Universiteitsbibliotheek Uithof, but two interviews were held in train station coffee shops.

Based on the interviews and screening survey (Appendix B) several demographic and psychological factors can be disclosed about the participants. The sample consisted of three men and six women aged 18 and 33 (mean 24.0, SD 4.5). All participants were born in and had spent most of their lives in different countries; four were from countries in Europe, three were from Asia, one was from North America and one was from South America. One had grown up in the Netherlands and two spoke Dutch. Most, but not all, participants were in or had completed some higher education in the Netherlands. Five participants had a C2 level of English fluency, two were native English speakers, one had a C1 fluency level, and one did not know their fluency level. Based on the screening survey and the interview, five of the participants had been diagnosed or suspected to have one or more of the following diagnoses: ADHD, autism, depression, anxiety, and eating disorders.

### 5.3.3 Psychometric Data

The psychometric data were gathered from the screening and in-person psychometric surveys (Appendix E). This yielded two measurements of participants’ loneliness with the UCLA-LS-9 psychometric, and one score for each of the remaining psychometrics. These scores can quantify psychological phenomena and reveal opinions and experiences that were not otherwise shared in the interviews. All psychometric scores for the participants are available in Appendix I.

Between the two UCLA-LS-9 results, each participant’s average score was between 20.5 (64th percentile) and 26 (90th percentile). The lowest recorded

UCLA-LS-9 score was 18 (48th percentile) and the highest was 28 (94th percentile); these were recorded during the interview's psychometric survey after both had screening scores of 23. The test-retest validity, indicated by the Pearson Correlation Coefficient between the two tests, was low at 0.1678.

For the psychometrics addressing depression and generalized anxiety, the PHQ-9 and GAD-7, the guidelines associate qualitative meanings with the scores. For the PHQ-9, 1-4 is minimal, 5-9 is mild, 10-14 is moderate, 15-19 is moderately severe, and anything higher than 20 is severe. The scale of GAD-7 is similar for 1-15, but anything higher than 15 is classified as severe. Participant's PHQ-9 scores ranged between 3 (minimal) to 17 (moderately severe), with an average (SD) of 8.444 (4.186), and their GAD-7 scores ranged between 2 (minimal) to 12 (moderate), with an average (SD) of 7.222 (3.4197).

The psychometrics for social anxiety and social phobia have individual cut-offs, but the severity is judged simply by the percentile associated with the score. For the SIAS-6, five participants scored below 7, indicating no social anxiety. Three participants had medium social anxiety at the 31st percentile, and one had a score of 21, connoting high social anxiety at the 95th percentile. For the SPS-6, three participants scored at or below 2, indicating no social phobia. Four participants had low social phobia, indicated by scores below the 25th percentile, one hit the 68th percentile, and the last participant reached the 96th percentile.

When measuring participants' positive and negative affect with PANAS, interpretation again relies on converting the scores to percentiles. Two participants had positive affect at the 82nd and 91st percentiles, but the rest placed below the 20th percentile. All participants had negative affect scores above the 71st percentile, and five participants had a negative affect score above the 90th percentile.

#### **5.3.4 Common codes and requirements**

The interviews covered diverse topics, and each participant brought unique and varied experiences and opinions to the discussion. These results cover the most common codes and opinions expressed by at least five participants; some less common experiences will be highlighted in Sections 5.4.1 and Section 5.4.3.

**Experience of loneliness.** Most participants spoke about how they felt different than the people around them. Some pointed to differences in language or nationality, like when I2 shared that "*When I did my Bachelor's, for example, we were 200 people and I was the only one who wasn't [nationality], which wasn't very nice.*" Others had differences in social behaviors, intrinsic identities, and interests; I9 explained that "*I've never met anyone who experienced their sexuality the same as I do.*" Additionally, several participants mentioned how social comparison and their interpretations of social norms also impacted their experience of loneliness. I3 talked about how "*I try to do things because I saw people doing*

*things like that and I thought that's the norm, I should do things like that as well. I should have a partner to do whatever things together, or good friends to do things together."* I2 echos this idea: *"I see other people succeeding very easily, but not myself. So I've been thinking, what's wrong with me? There's something wrong. Because everyone else managed to do it."*

Participants mentioned that elements of their mental health have impacted their loneliness. I5 mentioned how their anxiety plays a role, *"I know that a lot of the feelings are based on me kind of, like, catastrophizing [...] It's a lot of the results of the anxiety."* I8 went so far as to suppose that *"In general, all my loneliness problems come from my eating disorder."* Participants talked about the difficulties that their environments posed when trying to make or maintain friendships. I1 referenced their social and international experience in their minor, arguing that their loneliness was *"mainly due to the environment that I'm in. Like if I'm in an environment like [during] the minor, I would not say that I feel lonely."*

Most of the participants had experienced changes and improvements in their experience of loneliness (e.g., *"It has changed a bit, I think, since coming to Utrecht, it's got a bit better."* - I2), their attitude toward loneliness (e.g., *"I feel like I'm not so afraid of loneliness anymore."* - I3), and their ability to deal with loneliness (e.g., *"It helped me develop a lot of coping mechanisms so that I can be fine whenever I'm on my own."* - I4). However, they had also experienced times when increased solitude also led to more loneliness; several participants shared times from the COVID pandemic and ensuing lockdowns that affected their loneliness.

All participants talked about people in their social network, with varying levels of apparent appreciation and satisfaction with these relationships. For example, I9 shared that they *"now have the friends around me where I can be who I am. Which means I'm starting to show people more and more myself,"* I3 mentioned *"I'm well connected to my family, and I have a real warm family,"* while I2 said that *"I do have a couple of friends in [home country], but you know... Yeah, it's not that much, really."* Most of the participants talked to their close social network about their experiences of loneliness, though several emphatically stated otherwise. Some leaned on their friends when feeling lonely, as explained by I9: *"What I try to accomplish with talking to friends is to make sure that the peak [of loneliness] never happens."* However, some participants were only interested in talking to their social network after they felt better, like I3, who clarified that *"When I've healed myself. I go out and tell you, like, "Oh, I feel lonely last week, oh, I felt terrible, I cried." But I was by myself."*

**Helpful and unhelpful conversational elements.** Based on the interviews, the most important element that's needed for a helpful conversation about loneliness is the simple presence and support of the conversation partner. I3 empha-



sized this as the most helpful element for them: *“I don’t need solutions from them, I don’t need encouragement from them. Just someone who’s there.”* And I9 explained that, even if the conversation partner is unable to offer advice or other forms of support, *“at least they’re willing to help, and that feeling already improves the situation.”*

Another helpful conversational element was reassurance and reminders of the lonely individual’s social network, intrinsic value, and the validity of their emotions. I3 shared a time when they talked with their parents about their struggles with loneliness; their parents offered the sentiment that *“this is who you are, and what you are doing is good. If you think it’s fine, it’s fine. You don’t have to be the same as others.”*

Another key group of helpful conversation strategies is supporting the lonely individual to think more about and identify the key elements they’re struggling with. Participants found value in speaking the words into existence; I3 posited that *“You might not be so clear about your thoughts. You just keep talking and they make us make a conclusion.”* Some participants believed that specifically asking about the sources of thoughts and the motivations behind actions could be helpful. I4 brought up the value of questions *“that help you introspect, that you may not be able to ask yourself at the moment, but then, if someone else asks you, then you start thinking in that direction. That might help.”* Finally, simply allowing a distressed person space to vent was suggested as a helpful offer. I4 talked about how *“Sometimes just talking about it also helps, just acknowledging it in front of somewhere else that I’m feeling this and this and it’s not entirely in your head. And then you stop spiraling.”*

When asked about harmful or unhelpful conversational components, participants had more varied responses. The most common responses focused on a lack of empathy, as represented by I2’s example: *“If there’s a problem, “Oh, just deal with it”. You know, that doesn’t really help, it just internalizes the loneliness more.”* In addition, people brought up instances where they felt others dismissed their experiences of loneliness, as shared by I8 about their parents *“They’re going to maybe judge me, or saying “Oh, you are not grateful at all because you have a house where you live, you have food.””*

**Opinions around technological mental health interventions.** One of the common things that participants mentioned when talking about appealing elements of conversational AI interactions was availability. I9 simply states *“Unlike a person, you know, an AI tends to be available whenever.”* This was also mentioned by more hesitant participants like I5, who said, *“I’m sure if I got lonely enough, I’d be open to talking to the computer, knowing that it would at least respond to me.”*

However, limitations of CA interactions were also discussed by most par-

ticipants, including the CA's inhumanity and inability to truly understand and empathize with the user. The reactions to this fact ranged from amusement, as exemplified by I9's statement, *"If the computer would tell me that, you know, they understand, I would probably laugh,"* to I7's distress, *"You might freak out that, 'Fuck, I'm talking with a robot.' You know, you might freak out with that idea."* I6 also explained their perspective that, even if the AI offers the same advice as a human would, *"with an AI, it's most likely that the persons [sic] are looking down on the AI than when you're talking to an actual human."*

**Reactions to the prototypes.** Once the participants were provided with the prototypes, they shared their feedback on the features and other interaction qualities. The thought analysis was intended to identify thoughts and assumptions, classify them by sentiment, and explain the classification. Participants thought this was impressive, and helped them trust the CA's ability to interpret their situation. I9 explained that *"Because the analysis is correct, I have more trust in what it then, what the rest of it says."* I6 added some nuance to this perspective, sharing *"I don't think people like to feel analyzed. But it is very interesting for me to see that [the CA] can, like, do it."* Additionally, participants believed the thought identification helped them break down assumptions and make their thoughts clearer. I5 stated that *"Having your thought patterns kind of presented to you, that was really helpful."* I8 expands on this, *"When they explain it more rationally, you start to understand what you are feeling and then you stop, for example, being overwhelmed and it's like, 'Ok, this is what is happening right now.'"*

The next contribution by the CA was the perspective shift, designed to use the identified thoughts and craft new ways of thinking about the situation. Some of the opinions participants offered later conflicted with their own initial thoughts. For example, they shared that the perspective shift was too general, or that *"it's not suggesting things I haven't thought of myself already,"* as explained by I9. They shared that, in the moment, the perspective shift may seem unrealistic (*"This seems like an ok conclusion, right? Yeah. But maybe they need two years to realize that."* - I3), or, as I5 said, *"The perspective shift advice is, like, good advice in general. I feel like, in practice, it might be difficult to implement."* They added that the CAs did not seem to know enough about their situation to offer personalized, useful advice in the perspective shift. I5 suggested that *"if it could continue the conversation in that direction and ask me that, 'What is something you like to do, is something you generally like,' and then push me in that direction. So I don't feel as much inertia in picking some other activity that could help me with my loneliness."* Especially when comparing the two prototypes, participants identified that the chat's ability to ask further questions was important, as explained by I7: *"It maybe needed more information, so it would need to ask you for that information to understand the situation. [...] it needed to be more interactive, to have one more*

*round of input that you give it.”*

However, some participants remarked that the perspective shifts offered useful ideas that would be helpful in the moment. I8 explained *“I feel like, if I could tell someone and [they could] give me these type of rational thoughts, I could see everything more clearly.”* Similarly, I1 said *“if you’re struggling with negative emotions or, like, negative thoughts, it definitely gives you a positive perspective and how you can, like, also think positively. Like, maybe you’re not in the best space right now, but you could still help yourself by thinking positively.”* If used in a true situation, I2 said *“I would try to implement it. You know, if it’s asking you to reinterpret the situation, I would probably try to do that.”* I7 acknowledged both the generic nature of the analysis and the help it can offer by saying *“I don’t think it says something that I didn’t think of, also. But even just saying something that you’ve already thought is a validation, in a way.”*

The feedback about the prototypes was relatively mixed, but one of the other common pieces of feedback related to the chat-based interaction. Participants pointed out elements of the interaction that were clearly artificial or reminded them of other chatbots. After explaining a preference for the chat, I6 clarified that *“It’s kind of broken, you know. Like, because it’s a chat, I feel like I’m chatting to someone, but then also, it talks like an AI.”* I4 pointed to an indicator of artificiality and offered a suggestion: *“For everything I say, it validates it, but then it validates every sentence of it. Like, “Ok, I can understand you’re feeling lonely, I can understand your mood, I can understand...” But you know, that makes you feel like this is... it doesn’t feel like a human connection, so that can just drive someone away from it. So, if it’s just shorter, like “I understand you feel this way,” then just go ahead and not validate every sentence that I wrote, that might be better.”* However, they also shared that the interaction was easy and understandable. For the form interaction, I4 stated *“There are certain questions, you fill them in, and then it analyzes and tells you something. So that expectation was met.”* And I1, when describing the chat interaction, said *“if you already know how ChatGPT works, it’s kind of the same. And, yeah, it’s very easy. You just have to write and then put enter, and the computer, or the AI, does the work itself. So, yeah, it’s easy to use.”*

**Expectations of a conversational AI interaction.** Finally, participants discussed the experience and the elements of the interaction that they would expect, based on their preconceptions and their experience with the prototypes. One of the uncontested opinions, though not spoken by all participants, was the appeal of the system offering psychological and research-backed strategies and other mental health resources to users. For example, I1 explained that *“I would have liked to know more about like theoretical stuff [...] like what maybe psychologists themselves use to help people with those kinds of feelings.”* I3 added *“maybe you could*

*have like a partnership with online therapy or something.”*

A recurring theme throughout the interviews was the focus on and importance of data privacy for CA interactions about mental health. The proposed methods of achieving this varied from anonymity (e.g., *“I’m assuming that I’m anonymously talking to the computer and that it’s not storing what I say or anything.”* - I5), to offline, hosted interactions (e.g., *“If it’s completely offline, just whatever storage is there is on my phone, and not being stored or being shared.”* - I4), to simple assurances of safety. I2 explained that the problem was specific to this application of AI: *“I have no problem using AI for other things, but if I’m talking about, like, yeah, mental health, problems and stuff, and personal things. It would be really difficult.”*

However, most participants shared that once they were unconcerned about their data security, they would want to talk with the CA about any and all of their issues. I8 explained that there would not be any topics that are too private because *“I don’t think I will be, feel invaded if I’m talking with an AI because it’s not a real person.”* I3 had a similar sentiment, *“Actually it’s more important if they ask personal questions because you get more intimate and you can say the things you really think about. Otherwise you just go to a friend, or, like, a random person. You go to an AI, you need some privacy, but the insight comes from the conversation. I wouldn’t want to hide anything or I wouldn’t want to reserve information for myself. I would like to be able to just talk.”*

Due to the hesitancy around data privacy and security, the idea that the system should remember the user’s previous conversations and other factors about the user was more varied. Still, most participants believed that the system’s efficacy depended on it being able to learn about the user’s situation over time as explained by I3: *“I think for getting over emotions or difficulties like loneliness, this is a long term thing.”* Some participants further focused on the CA’s ability to identify the user’s goals (*“I would expect AI is to get an idea of who I am, what is, what it is that I’m looking for.”* - I9) and personality (*“I would like to build and show what my personality is and then talk about it to make it make sense”* - I6).

Participants were prompted about whether it would be beneficial for the CA to have a discernible character, including a face and a name. While most participants were not opposed to the idea of the CA having a name or face, the main consensus was that the user should be able to adjust certain elements of the CA’s personality and appearance. I2 *“I know like the Bing chat, there are like three options. There’s more technical, clear, there’s a more creative one. So maybe what would help is having a different choice of personalities that you could have, and then that might be more tailored to the actual person.”* I8 offered a more gender-focused possibility, that *“you can choose between, for example, avatars, different avatars. And you can choose between these, if it’s a girl, it’s a boy.”* It

would be better, I4 explained, if the adaptation could happen automatically, *“if it could detect, understand. . . Let’s say I’m really depressed and then this character comes in and it’s giving all replies which are like very like enthusiastic, and very, you know, “Just do this! You got this!” and I would not want to listen to all that.”*

A final, more controversial, topic related to the interaction mode, mainly a comparison between a typed and a spoken interaction. Most participants preferred a typed interaction, often due to a sense of anxiety or nerves. I9 explained *“I like to think about my words and rewrite my sentence three times before I send it,”* and I5 expanded on this, *“It makes me really anxious because I’m like, “Oh my God, I can’t breathe, like, the microphone’s going to pick that up and get confused.””* However, due to some individuals having more nuanced statements, the majority also found value in spoken interactions. For example, I6 stated that *“I think I would rather text it, but I think it would be helpful for a lot more people if it were, like, a conversational interaction.”* Spoken interactions were seen as more emotive and helpful to the speaker because, as I4 said, *“You’re not just thinking in your head. You’re saying it out loud, so that has a different impact than just writing it, or just thinking it.”* Participants also said that a spoken interaction could make it feel more human-like, that, as I2 commented, *“If you’re just chatting, it just, it feels more artificial. At least if it’s a voice, like, it’s still artificial, but it’s less artificial-feeling, and then some people might not be bothered by that.”*

## **5.4 Analysis**

The interviews illuminate the desires and opinions offered by most participants. With further analysis, however, it’s possible to find patterns in participants that can be used to develop distinct groups. It will become clear that, although there are commonalities among participants, many of the opinions and experiences are too diverse to have one solution be the best option for all.

### **5.4.1 Codes**

Based on the common codes described in Section 5.3.4, feeling different than others, social norms and social comparisons, moving, one’s environment, and elements of their mental health are some of the top relevant factors to peoples’ experiences of loneliness. Participants mentioned having a social network and could recall times when their loneliness improved and times when it worsened. The most common helpful conversational elements related to the presence of someone willing to listen, reassurances of valued connections and their intrinsic value, and validation of their emotions. Prompting the lonely individual to express their emotions, examine their thoughts, and think critically about their situation were also common strategies. The most repeated harmful conversational elements were

a lack of empathy and the dismissal of the value of their perspective.

Participants appreciated the CA's availability but countered that its inhumanity, which prevented it from understanding the user, limited the interaction's value. The participants found the thought identification's accuracy impressive, thought it inspired trust, and enjoyed how it clarified their thoughts and identified assumptions. In some situations, the perspective shifts were sometimes too general, unrealistic, and not knowledgeable enough on the subject. However, some perspective shifts were assumed to be helpful in the moment and the chat's ability to ask questions helped make them more detailed and helpful.

Psychology-based interventions were regarded positively. Ensured data privacy and anonymity were valuable options to ease participant's worries and increase comfort in talking freely with the CA. Some participants devalued anonymity in favor of the ability of the interaction to remember them. Participants were not drawn to or opposed to the CA having a name and a face, but the appeal of customizability arose as a middle ground. Participants argued for both text-based interactions, described as easy to use and less anxiety-inducing, and voice-based interactions, which were seen as more emotive and human-like.

However, we can also learn from the characteristics, including gender, anxiety, depression, and affect levels, that divide participants. The group of participants with strong social anxiety is the same as Profile E, covered in the profile-based analysis in Section 5.4.3. If a characteristic is described within a group below, it is most connected to or is unique to this group. It is important to note that a participant not talking about a certain topic is not valuable insight into their perspective and the trends can only be built on the presence, not the absence, of conversation.

**Men:** Three men participated in the interviews (I2, I6, and I9); these participants shared very little in terms of psychometrics, but offered some distinct opinions. While over half of the participants mentioned searching for people who were incredibly close in their understanding about the world, the only two participants who talked about wanting to be in romantic relationships with these people were the men. Additionally, all members of this group and only one woman talked about how simply being in the company of others was not enough to abate the loneliness, as I2 pointed out that "*it's less to do with, like, being with people and more to do with the quality of being with them.*" Similarly in the distribution of this opinion, the entire group mentioned that various societal expectations about who they should be (including non-emotive, tough, and individualistic) increased their loneliness. While none of the helpful statements were distinct to the men, offering unactionable and unempathetic platitudes like "*Man up,*" or "*Have you tried thinking positively about this?*" was mentioned by two of the men.

While hesitancy around the interaction's application was more widespread,

two of the three people who mentioned a bias against AI were men, including I6 sharing that *“it’s most likely that the persons are looking down on the AI than when you’re talking to an actual human.”* All the men also found the thought identification useless, often referencing that the information was already known. Instead, it seemed that using the chat to talk through the issue was the most important element to two of the men and only one woman and that the analysis at the end was less helpful, as explained by I9, who posited that *“in a lot of cases, you can come up with the answers yourself. Because in a lot of cases, you know the answers.”* The only participants who did not care about what happens to their data (*“I don’t really care about private information.”* - I6) and the only two who mentioned the uncanny valley or that the CA should not try to trick the user into thinking its a human were men (*“there needs to be like, OK, this is not real, this is not human.”* - I2).

**Women:** The remaining six participants (I1, I3, I4, I5, I7, and I8) were women and were not different psychometrically from the men. Four of the five people who accepted their solitude or loneliness were women, like I3 who shared how *“if you accept it, if you adapt to it, it can be something good for your life as well.”* Five of the six women were the only people who talked about prioritizing things other than socializing. I7 explained how their field of study *“helped me to fight that loneliness because I was feeling like I can spend my time with it, and develop my skills, and this is what matters.”* All the women and only one man talked about how spending more time on their own increased their loneliness, three of whom mentioned the COVID pandemic and subsequent lockdowns. Most of the women in the sample talked to their social network about their loneliness, but four of them avoid talking about it when they’re feeling their lowest.

The women involved in this research offered many conversational strategies that were never or rarely mentioned by the men. In addition to be willing to listen to someone and offer empathy, five of the women and only one man mentioned offering special attention, *“telling me “Oh, whenever this happens again, just call me, and tell me to meet, and stuff,””* which helped I7, or making plans to *“make plans and try to make them go out,”* as mentioned by I8. Four of the women and one man mentioned the benefit of challenging your ideas and seeing things from another point of view, which can help answer questions like I4 asked: *“Am I really lonely, or is it something else which is bothering me?”* While men mentioned the value of being able to vent to the CA, four of the five people who mentioned doing this in conversations were women. The participants who valued the conversation partner’s ability to relate to their experience, those who appreciated it when people offered strategies that worked for them, and those who promoted self-management were primarily women. Additionally, the only three people to mention getting out of their comfort zone were women, including I3 who recounted how they took

their friend's advice *"to invite friends over for dinner, even though I didn't feel comfortable at all, and I was like "Ok, I'm going to throw a party." And I'm like, "I'm throwing a party?!?" But then it turned out good, and it's about learning from other people."* When it came to unhelpful statements, only women mentioned that it would be unhelpful to be mean or aggressive, to theorize about someone's mental health, or assume that they are fine when, as I5 suggested, *"maybe I would be going to the chatbot to ask for coping advice."*

The only three people who mentioned that sharing their emotions was easier with the CA than with people were women, who referenced the AI's inability to judge you or gossip about you. Half of the women believed that the thought identification itself validated their interpretations of the situation, or, as I5 said *"I guess it can just be nice to read them and, like, get that kind of external reinforcement."* Similarly, half of the women believed that the thought classification, which pointed out if a thought was positive or negative, reinforced negative feelings; I3 shared the *"I would not want to read this because I already know they are all negative."* The four people who mentioned limiting the length and the number of questions per message were also all women, who felt overwhelmed by the amount of information it would send at a time. It also seemed that only the women would have appreciated more direction on how to best communicate with the interactions. When it came to future use, only four women mentioned the need for the CA to have safeguards in place for extreme mental health concerns or *"if, let's say, someone is suicidal, it should be able to tell you that, "OK, you are in this such a state and you should consult someone, not just the AI,""* as mentioned by I4. Finally, only women mentioned data security solutions or concerns; two wanted to know how their data would be secured, three wanted the interaction to be anonymous, and four wanted the interactions to be immediately deleted.

**Depressed People:** Four participants (I1, I2, I5, and I7) had clinical levels of depression (above the 17th percentile). Three of these participants connected positive changes to their loneliness with changes in the environment, including their country and social circle. The only two people who felt unwelcome in their community and noticed that this affected their loneliness were in this group, as explained by I2's assertion that, *"if you don't identify as [Country Demonym], then you're not really part of the group."* Three of the four people who identify social or cultural barriers affecting their loneliness were also in this group. Two of the three self-described introverts and the only two people who identified an imbalance of effort in their relationships were depressed. Further, three of the four people who discussed how rejection and other tumultuous social dynamics impacted their loneliness were a part of this group, including I2's assertion that *"I've been part of groups in my life, but there's always been something that's happened that has sort of ruined it,"* and I5's experience of *"We had another*



*argument this fall [...] now our friend groups don't hang out much anymore."* These participants often compared their social lives with those of others and their internalized social expectations, exemplified by I2's statement that *"I see other people succeeding very easily, but not myself. So I've been thinking, what's wrong with me?"*

The only three participants who said that they found the analysis given at the end of the chat helpful were in this group; I1 explained that *"at the end, like, when you're done chatting, it gives you that analysis; it's nice. [...] So you can read it afterwards."* Additionally, this group also included the only three participants to agree with the idea that it would be beneficial for the CA to offer summaries of past topics discussed.

**Pessimistic, Anxious Women:** In the sample of participants interviewed in this study, three women (I1, I5, and I8) were the only ones with clinical levels of GAD (above the 65th percentile) and also had the highest levels of negative affect (above the 97th percentile). Two of these participants overlap with the previous group, resulting in some repetition, and all of these participants are women, so these statements are also primarily relevant to the women participants. Two of these participants claimed to do nothing for their loneliness; I1 said they *"I don't really have any other methods or things that I do to help that feeling,"* and I8 shared that *"I wouldn't deal with loneliness, like, I will just let loneliness eat me."* All of this group and one other woman talked about feeling stuck in their environment, or that *"I think moving helped me a lot,"* as mentioned by I8, and that socializing can take more effort than they tend to give, as exemplified by I5's statement that *"I'm sure that, if I really tried, I could have been more involved, or reached out."*

Two of these participants and one other woman found it helpful for someone to help summarize the issue to clarify their thought process; I8 wished for the ability to *"tell someone and [they] give me these type of rational thoughts, and I could see everything more clearly."* All of and only these participants mentioned that blaming someone for their situation was not helpful because, in I1's words, *"it's not their fault, most of the times, that they feel like that."* In reaction to the prototype interactions, two of these participants found the analysis of the chat to be a helpful addition as it gave them more feedback while still getting *"the feeling of kind of talking to someone,"* according to I5. Additionally, while several participants outside of this group mentioned the usefulness of the ideas offered by the perspective shift, the members of this group were most enthusiastic about it; I8 shared that, *"in that moment, I wouldn't think, "I can do anything else." I was really thinking "I can't do anything else. I will probably lie in bed all day, disappointed." So, I think that's really nice."*

**People without Depression or Anxiety:** Finally, three participants (I3, I6,

and I9) had sub-clinical levels of both depression (below the 8th percentile) and GAD (below the 44th percentile). Two of these participants were assigned to profile A, and commonalities between I3 and I9 are expanded upon in Section 5.4.3. Two of these participants, as well as two others, found simply spending time socializing could improve their loneliness, perhaps as a distraction as mentioned by I6: *“I’m just trying to not be alone most of the time because I have friends and I spend time with them, maybe more now than before, to maybe not get stuck in my own thoughts.”* They also found it, as I9 explained, detrimental for people to take their problems too seriously, that *“once they make it theirs, they blow it up to larger than it actually has to be. And then they are becoming far more in emergency mode than is necessary.”* All three participants found it helpful to use the act of talking as a way to unravel their thoughts, *“like people going to church to confess,”* compared I3, and I9 added that *“if you have to explain to someone else what you’re feeling, why, what, etc., it paints a far clearer picture for yourself.”*

All three participants did not find the thought identification useful, considering it unnecessary repetition, and were critical about some of the CA’s phrasing, which they found patronizing or disconnected them from the message. For example, I6 commented that *“I feel like with this just shoving into your face, it’s like ‘I know you’re lonely. Just tell me why,’”* and I3 added that *“The subject is ‘I’, so it’s from the user’s perspective. [...] when I read this, after so long, I feel like ‘Oh yeah, this is how I feel’, but after the process. [...] but if I am in the process, I will feel disconnected with the ‘I’ there because I’m not there yet.”* One of the participants agreed with I3 that *“I would want it to have a face, I would want it to have a voice, not just texting.”*

#### **5.4.2 Profiles**

When comparing the psychometric data from the participants, several patterns were found that enabled them to be grouped into profiles. The use of profiles for loneliness interventions was inspired by the paper by Käll et al.<sup>8</sup>, in which they used PHQ-9, GAD-7, SIAS, and UCLS-3 to identify latent psychopathic profiles. Käll’s goal with the profiles was to describe the population seeking help for loneliness; they used the psychometrics of around 330 participants.

Unlike Käll’s profiles, however, this research utilized a different version of UCLA-LS and SIAS and also used the SPS-6 and PANAS psychometrics. Further, Käll’s main mode of discerning profiles was latent profile analysis and other statistical tools that would not work with a significantly smaller sample size. Additionally, with only 9 participants, this research does not aim to describe the entire population; the aim of these profiles is instead to find patterns in users that can inform the personalization of interventions. Finally, the profiles do not describe participants’ relation to measured loneliness because the population was already

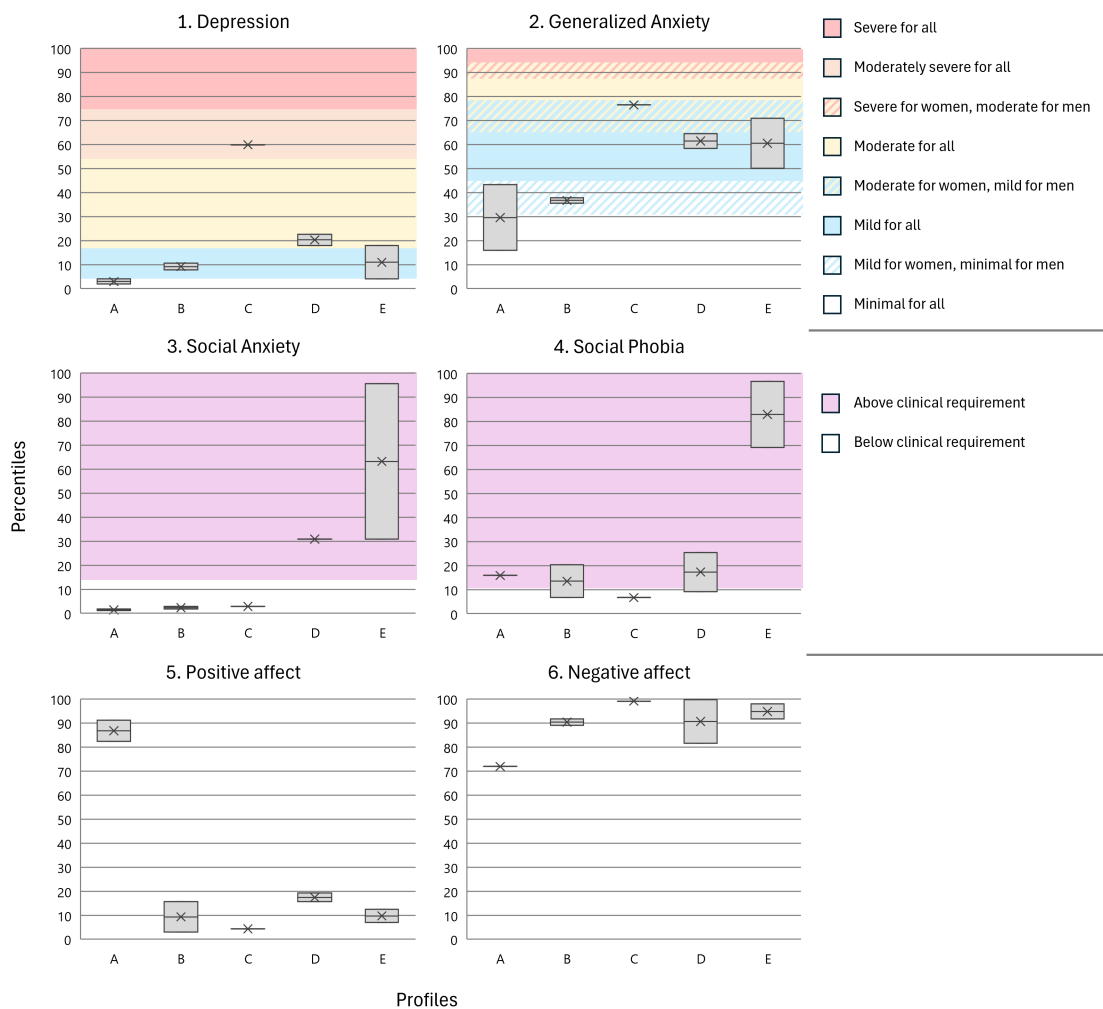
selected on this criteria and was not a discernible quality of the participants' profiles.

To identify the profiles, the psychometric scores and the percentiles for each of those scores was used. The scores alone were used for the identification of psychometric tendencies. For example, based on the literature related to the PHQ-9, a score above 10 indicates a diagnosis of depression, which grouped I1, I2, I5, and I7. The percentiles of the scores were used to contextualize the psychometrics that did not have set meanings for each value and also to stratify the scores further. For example, the SPS-6 specifies that a score above 2 indicates social phobia, but there is a large difference in a low score like I3's 4 and I8's 21, and this is more clearly reflected in their percentiles, 15th and 96th, respectively.

From these and other patterns, profile A was clearly distinct due to the high positive affect, lack of depression and anxiety, and the lowest negative affect of the sample. Profile E was also clearly the most socially anxious and socially phobic individuals, Profile C included the two participants with the highest depression and no social phobia, profile B had very similar scores across all but a 15 percentile difference in the social phobia psychometric, and profile D was different due to their relatively high scores across the board. An overview of all the profiles is available in Table 2, and the psychometric qualities can be visually compared in Figure 7.

**Table 2.** Profiles identified from the psychometric data. \* indicates more variability between the profile's members ( $\geq 10$  percentiles).

<b>Profile (IDs)</b>	<b>Depression (PHQ-9)</b>	<b>Generalized Anxiety Disorder (GAD-7)</b>	<b>Social Anxiety (SIAS-6)</b>	<b>Social Phobia (SPS-6)</b>	<b>Positive Affect (PANAS)</b>	<b>Negative Affect (PANAS)</b>
A (I3, I9)	Minimal	Minimal*	-	Low Severity	High Levels	Medium - High Levels
B (I4, I6)	Mild	Mild	-	Low Severity*	Low Levels*	High Levels
C (I1)	Moderately Severe	Moderate	-	-	Low Levels	High Levels
D (I5, I7)	Moderate	Mild - Moderate	Medium Severity	Low Severity*	Low - Medium Levels	High Levels*
E (I2, I8)	Mild*	Mild*	Medium - High Severity*	High Severity*	Low Levels	High Levels



**Figure 7.** The psychometric ranges for the profiles including all participants. The hatched areas of Chart 2 account for the gender-based population differences; men had lower average GAD scores, so the same cut-off between severity levels translates to a higher percentile for men than for women. Profiles A, B, and E each contain one man.

When comparing the profiles identified in these interviews with the ones in Käll’s research, profile A seems to align with his profile 1 “Mild psychopathology” due to the below-average levels of depression, GAD, and social anxiety. Profile B also aligns with his profile 2 “Moderate depression and worry” due to the low social anxiety but present depression and GAD. Profile E matched his profile 3 “Primarily socially anxious”, as they both had high levels of social anxiety and mild to moderate levels of depression and GAD. Profile D is also similar to his profile 5 “Moderate psychopathology”, which has clinical social anxiety and moderate depression and GAD symptomology. His last and rarest profile, known

as “Severe psychopathology”, does not match any of the participants; profile C has the highest depression and GAD scores of this sample, but is missing the strong social anxiety present in his profile 4. However, since this research also captured data related to social phobia and positive and negative affect, these profiles can be more detailed.

### 5.4.3 Profile-based Code Analysis

The participants in each profile shared many opinions and perspectives, as explained in Section 5.3.4, but each profile offered additional thoughts that were unique or more strongly held. Highlighting these differences can inform how the psychometric similarity may manifest.

**Profile A.** This profile consists of I3 and I9, who felt strongly that they were different from the people around them, but had also accepted this difference as part of their reality. I9 explained that *“It was just very clear that the way I experienced the world was not the way most other people experience the world.”* They were aware of their emotions and willing to talk about them but were worried about the effect of their emotions on others. I3 mentioned that *“when I’m angry, when I’m feeling sad, feeling lonely, I don’t want other people to notice. Because I feel it’s not their responsibility.”* I9, after mentioning how they rely on conversations with friends to help with loneliness, clarifies that *“it doesn’t make it their problem. And that’s also what I like, because I don’t want to burden them, necessarily, with my problems.”* They believed that helpful conversations about loneliness should include reassurances of how it’s fine to be who you are and that your relationships are stable and expressions of empathy and care. Additionally, they found it helpful for their conversation partner to be optimistic but realistic, relate to their experiences, and to encourage self-management. The conversation should also allow space for them to talk through and unravel their thoughts without jumping to advice or making assumptions. Participants talked about how important it was to not feel like their experiences were minimized or taken too seriously, and they disliked unempathetic, unrealistic advice.

In reaction to their interactions with the CA, these participants appreciated the 24/7 accessibility but believed that much of the point of talking about their emotions was to bond or connect with people, making the AI interaction less worthwhile. While the thought identification helped show them that the CA was working with an accurate interpretation of the situation, increasing trust, it was also obvious and not helpful. The perspective shifts were also too general and occasionally unrealistic, though they also mentioned statements from this section that were ideas that they had grown to accept.

They believed that the interaction’s goal should be to help them develop small, actionable steps toward their goals, and thus should remember them and

their context to help make it more personal. In the interactions with the CA, these participants had a balanced view of the impact of their actions and agency, used approximation words like “seemed to” and “as if”, and focused on listing feelings and worries experienced in the moment.

**Profile B.** These participants, I4 and I6, attributed a large part of their loneliness to their environment, including moving to another country. They had a reliable group of friends and family but struggled to feel connected deeply to people. This meant that, when they moved halfway around the world away from their existing social network, they had a hard time developing new, deep social connections. They seemed to be in tune with their loneliness, attempting to diagnose their emotions and the causes behind what they believed to be a common experience with their peers. They used this experience and awareness to offer counsel and advice to their friends dealing with loneliness. They also avoided talking about their own loneliness with their loved ones because they didn’t want to burden them with their emotions.

In conversations about loneliness, they found it helpful to be reminded that their emotional reactions are valid and that their conversation partner was empathetic and willing to listen. It was helpful to offer an outside, more objective perspective that could help pull them out of their emotions and process the situation more rationally. Questioning the reasons behind their thoughts and actions helped them think more deeply about their motivations, and allowing them space to talk freely helped them get their thoughts out of their head and unravel their thoughts. They also suggested that simply being around people and having unrelated conversations was helpful in dealing with feelings of loneliness. They found it unhelpful to offer unprompted advice and comments and to feel forced into talking about their feelings and experiences. However, they both mentioned that what is beneficial or detrimental depends on the person and the situation.

They initially had low belief in the ability of a CA to help with loneliness, but by the end of the interview, shared that they thought that CAs like the prototype could help many people. They found the thought identification to increase their trust in the CA’s ability to understand and found it helpful to break down their thoughts and assumptions. Additionally, they identified several suggestions from the CA’s perspective shift that they found helpful or insightful. However, they believed the thought classification would reinforce the user’s negative feelings. They found the form’s output matched their expectations, thus was more appealing, but found the chatbot easy to use. The use of probing questions aided their ability to share about their experience. As reflected in their classification of what was helpful, they stated that just talking about loneliness was more important than the analysis at the end of the prototype.

Their ideal CA would be nice and empathetic, have a name and a face, and

feel like you were talking to a real person. They suggested having the interaction include multiple-choice survey questions, similar to the psychometric questionnaire, about the user to inform the CA about themselves. These participants were interested in learning more about how their brain and thought processes worked. They found it difficult to trust an AI with mental health information and wanted to be able to skip or change questions and topics about sensitive subjects. They posited that a spoken interaction would be better for the emotive quality of the interaction than a typed one.

In their interaction with the prototypes, they attributed their loneliness to their international status and identified other concrete struggles they encountered. They mentioned specific emotions and desires that were embedded within their thoughts. They expressed concern about certain elements of their loneliness, like the chronicity and the quality of their relationships from back home. They explored different ways of thinking about the situation and asked for suggestions on how to better handle similar situations in the future.

**Profile C.** This profile was only associated with one participant, I1. This participant did not actively try to dispel their loneliness feelings, partially because they were emphasizing the completion of their academic work over their social life. They expressed that their physical and social environment was a large factor in their loneliness, and talked about a time when they were surrounded by international students as a period that *“gave me an alive feeling again, like, being social again and meeting so many people and bonding with them.”* This participant attributed part of the impact of the environmental shift to them feeling different from Dutch people, even though *“I already speak the language and stuff, I just think that I don’t match very well with their vibe.”* The other environmental elements focused on their past lack of knowledge of Dutch, their current city of residence, *“the worst city ever in the Netherlands for young people,”* and the delay in their studies that has prevented them from their desire to move abroad. They explained *“As soon as I’m done with my studies, I feel like I’m gonna start living again.”* This participant identified as an ambivert or an introvert who likes meeting people. They shared that they have not mentioned their loneliness to their friends or parents but that *“they can, I guess, assume that I’m lonely without even saying it. Because it’s obvious. Like, if you’re not doing fun stuff with other people, then you’re going to feel that way.”* They also described themselves as *“not, like, an open person in general,”* and believed that *“people always, like, gossip and stuff. That’s [...] maybe a reason why I don’t open up about my feelings to people.”*

This participant thought that expressing their empathy and willingness to listen was important in conversations about loneliness, that normalizing peoples’ experiences was crucial. The participant talked about how helpful it was for them to be pushed out of their comfort zone and found it helpful when their thera-



pist recommended self-soothing and enjoyable activities that distracted them from their loneliness. However, aligning with their avoidance of emotional discussion, they believed that actions help them manage their loneliness more than words could and that being forced to share their feelings would be unhelpful. This participant also seemed to associate feelings of blame with loneliness, sharing how saying the *“things that, like, that person would say, blaming them or making them feel worse, that would definitely not help.”* They also mentioned that being *“mean or, like, hostile towards them,”* could be detrimental and cautioned against being *“rude or whatever towards someone feeling that way; you never know what that could lead to eventually.”*

I1 theorized that *“if it’s difficult for you to, like, open up to someone, I guess it would be easier to do it with AI.”* However, just as they had mentioned that actions matter more than words, they also said that, in some situations, *“you really need someone like, in person, to meet with.”* Still, they saw no downsides to trying the interaction, saying that *“if it helps you, why not?”* They mentioned that the interaction being free and having an endorsement from mental health professionals and other authority figures would prompt them to use it. They also suggested educating users about psychological principles and offering tools that psychologists would recommend. In addition to psychological advice, this participant also mentioned a desire for the CA to, more generally, *“give, like, an appropriate solution to, like, your problems.”*

The participant mentioned an initial conception of the system to include a survey that can direct users to personalized advice or allow them to *“compare how you score compared to, like, other people, or, like, the average.”* They did not have any strong reactions to the prospect of the CA having a name or appearance, or being able to customize those elements. The participant also preferred an anonymous interaction where they would not have to register their identifiable information. In that case, though, they were not worried about talking about sensitive topics related to their mental health.

When it came to using the interaction, the participant was impressed by the thought identification and inspired trust in the CA’s interpretation of the situation. They received feedback from the CA that they found insightful and thought that the positive perspective shift would be helpful when in a negative mood. This opinion was retained for the chat interface; *“when you’re done chatting [...] it gives you, like, that analysis. It’s nice.”* Although they felt the responses sometimes felt robotic or like ChatGPT, they expressed that chat interactions were more of a human interaction. They found the interaction easy to use and preferred typing over speaking, but needed some direction when trying to avoid using personally identifiable information. In their interaction with the prototypes, their discussed situation heavily focused on their context and covered years of their life. Their

thoughts and assumptions revolved around their relation to their peers and how that affected their self-worth, and most of the actions and behaviors discussed were those of others instead of their own. Feelings discussed were nebulous (e.g., “tough” and “regret”), but not all of the discussion was negative, as the participant also included optimism and positive experiences.

**Profile D.** The participants within profile D, I5 and I7, both grew up in small communities that they were excited to leave. They have a few close friends, including their family members, though most of them were now in a different country than the participants. Both participants talked about moving as a difficult adjustment and described a sense of loss and sometimes guilt for leaving these individuals and for the relationships they had lost. They report to be picky about who they spend time with, preferring to find a good fit and matching values with people. I7 explained that *“not everyone is for everything. [...] Being around with, that’s ok, but you cannot share with them all your thoughts.”* Although they left the towns they were born in, they still felt a disconnect with their surrounding community, seemingly due to differences in social attitude and cultural interaction preferences, and want to move again. They acknowledge their role in making friends, but seem to be resistant to making the requisite effort. This resistance is likely related to relatively recent painful social experiences they both have had. I5 mentioned fights between groups of friends that lead to increased isolation, and I7 mentioned disclusion in social activities and a lack of effort by friends. They mentioned how periods of social isolation or withdrawal increased their feelings of loneliness. Both participants registered generalized and social anxiety, reflected in their statements about catastrophizing situations and having compulsions to check for social contact. They appreciated talking to close relations about their feelings and their loneliness but preferred to avoid conflict with those related to a painful situation.

In conversations about loneliness, a willingness to listen and a sense of support from their conversation partners was paramount. The participants found it valuable when people offer their time, attention, and care; as I5 suggested *“It comes to knowing that the person cares about you.”* Further, giving the person dealing with loneliness a safe social space to unproductively vent and talk about their experience was highly regarded. They also found it helpful to get an outside, more objective view of their situation to help identify unhelpful thought patterns and assumptions. Although one of the participants suggested offering emotion-based coping strategies and distraction, these participants did not value solution-oriented conversations and instead highlighted how unhelpful unprompted or un-informed feedback or advice can be. They also pointed out that it would be unhelpful to blame feelings of loneliness on someone’s past choices or mental health.

One of the benefits that these participants found in the CA interaction is the

ability to talk about their feelings without burdening or depending on their friends and family being awake and free to talk. They mentioned in their interactions with the prototype that they were happily surprised by the help and advice that the CA offered. However, the CA's lack of true understanding and empathy for the user also limited its power; I5 explained that *"I'd probably still want to talk to one of my friends or, like, one my parents, who knows me."* The participants posited that knowing more about how the CA works, how it develops responses, and how it is different from other AI tools would help increase trust in the interaction. Additionally, they suggested that it would be better if the CA acted like a therapist and was based on psychological research. The participants differed in how they wanted to prevent their data from falling into the wrong hands, but agreed that, once this was prevented, they did not mind talking about anything.

When trying the prototype interaction, these participants found that the identification of thoughts did not tell them anything new but, as I7 mentioned, *"saying something that you've already thought is validation."* They did find some of the ideas presented in the perspective shift too general or unrealistic; they tended to critique the ideas and find reasons why they would not work. They identified questions that the CA did not ask or elements that they did not share that left the CA partially uninformed about the situation. One of the participants did find some value in some of the suggestions, primarily related to how they were thinking about the situation. These participants found some of the responses repetitive and too long, and they requested more direction about what to discuss and how to talk about loneliness. While one did appreciate the form interaction, they both found the chat interaction to be more natural and elicit more information. In the interactions with the prototypes, they focused on rejection and related social restrictions and clarified the social dynamics of the situations. They described themselves and their interactions both as a bystander and a participant at fault. They made assumptions about others' closeness, identified patterns in others' behavior, and contrasted people's behaviors with others to those with the self. Interestingly, unlike what they talked about in the interview, they also asked for advice on how to cope with loneliness or relationships.

**Profile E.** Two participants, I2 and I8, were grouped into Profile E, distinct from the other profiles due to their higher levels of social anxiety and social phobia. These participants valued their close connections but found them stressful to maintain. They were lonelier earlier in their lives and they do not think they will ever leave that feeling behind, but noticed a positive change since they moved away from where they grew up. Like most participants, they felt different from the people around them, but this translated into a seeming deficiency in being able to connect deeply with others; I2 explained *"I find it difficult to connect with people as well."* They also described themselves as nervous, socially immature, quiet,

and very introverted, which they said negatively impacted their ability to adapt in new environments. While I2 tended to struggle with group situations and forming more lasting connections, I8 further found it hard to be in public, *“I will still not go in a crowded place, never.”* Both participants found it difficult to make new friends and disliked some of the social activities that people around them enjoyed, like going to bars and clubs. They also both found that their mental health impacted their loneliness, limiting their interest in their social life or their ability to maintain connections. Both participants mentioned that they would often prefer to spend time on their own instead of doing things with friends, and had most of their social time in school where it was required for them to attend. Neither participant shared their feelings of loneliness with anyone in their life, though I8 mentioned that *“I think this is the first time I [am] talking about it.”* but assumed that their solitude had communicated to their parents that they were lonely. When prompted to chat with the interaction about loneliness, I2 explained their hesitancy with *“It’s difficult to think because I’ve never been in this situation, really, before.”* I2 connected this lack of experience with their gender, explaining that *“boys don’t really like to talk about feelings or emotions with other boys.”* I8’s hesitancy seemed to stem more from their fear of being a burden on whomever they shared their feelings with.

These participants stated that it was helpful to reassure lonely people about their intrinsic value and the care present in their social network, that, according to I8, *“they have people who love them, and that they are actually not that alone.”* They also mentioned asking questions about their motivations to either highlight irregularities in their thinking or to help clarify the next steps would be helpful. It was also suggested that discussing strategies that allowed people to support or improve themselves was a valuable strategy. However, perhaps due to their limited experience talking about loneliness, they asserted that the most helpful thing someone could do was to invite a lonely person to do an activity or spend time with friends. I2 simply suggested *“just doing an activity with that person. I mean talking as well, of course, is important but I think, [...] when you’re doing activities with people, like traveling, or watching a film together, or playing a game, for me, that’s what helps a lot with loneliness.”* When it came to the unhelpful conversational elements, these participants focused on not giving people enough attention or trying to compare hurt feelings.

The participants of Profile E believed that a CA interaction like the one tested could be useful in clarifying thoughts and calming down emergent situations. However, they also expressed doubt that a CA could offer help in the same way that a human could; I2 explained that *“I feel, personally, it would make me more lonely because I know it’s not real.”* I8 added *“AIs are super smart, but I don’t know if they can have this type of understanding of human feelings.”* They

felt that it would be beneficial for the CA to offer personalized advice to combat loneliness, or, as I2 said, “*suggestions for what I can do in the in the real world.*” They imagined that the CA should be kind, empathetic, and feminine; I2 mentioned that “*Girls are more easier to talk to,*” and I8 said “*I wouldn’t be comfortable if I have to speak with a man about my things.*” They found the idea of customization of the CA appealing, suggesting a choice between avatars or personalities, that could be tailored to the user.

When they used the prototype interactions, they agreed that, if they used a present issue of theirs, the perspective shift could have been helpful. I8 theorized that “*I feel like, if I could tell someone and [they would] give me these type of rational thoughts, I could see everything more clearly.*” They found the AI analysis boxes at the end valuable information and felt the chat option had a more human-like feeling. Both participants also mentioned that the chat’s ability to ask questions was useful as it was easier and, as I8 said, “*it makes you go deeper in your thoughts.*” In their interactions with the prototypes, they tended to use abrupt and extreme words (“suddenly”, “very”, “no meaning at all”, “only”, and “can’t do anything else”). They also framed the situation around how others’ actions impacted their feelings and described how their emotions affected their self-worth.

#### **5.4.4 Connecting profiles and CBT strategies**

Based on the profile psychometrics, the experiences shared, the conversational elements thought to be the most helpful and unhelpful, the maintenance model of chronic loneliness, and the literature surrounding CBT and loneliness, further connections could be drawn. These connections helped indicate some of the CBT strategies that could be the most helpful to people who fit into a certain profile. These connections are outlined within Table 3.

For example, when looking at the maintenance model related to Profile B, it is evident that many of the conversational elements most helpful for them revolve around negative interpersonal appraisals. Negative interpersonal appraisals are evaluations and thoughts that assume that people one encounters are judging or otherwise thinking critically about them. Profile B’s focus on the reassurance of emotions and experiences, outside perspectives, aided exploration of motivations, and shared thought processes all contribute to counteracting or limiting the integration of harmful beliefs about what other people are thinking. A higher attention to others’ thoughts is not necessarily a detrimental tendency, but a lack of true understanding leads to assumptions, which can be damaging. To counteract this, Käll identifies several relevant strategies, including Cognitive restructuring, reviewing relationship experiences, psychoeducation, life review, and identifying personal strengths<sup>5</sup>. Cognitive restructuring, relationship experiences, and psychoeducation are the easiest strategies to implement with a CA due to the shorter-

**Table 3.** An outline of the elements of the maintenance model of chronic loneliness (outlined in Figure 2), the CBT strategies connected to them, and the profiles that best connect to these elements and are hypothesized to best connect to the associated CBT elements.

<b>Maintenance Model Element</b>	<b>CBT Strategies</b>	<b>Associated Profiles</b>
3. Mismatch between actual and desired social situations	Mindfulness	A
4. Negative interpersonal appraisals	Cognitive restructuring, Relationship Experiences	B, C, D
5. Counter-productive behavior	Exposure, Mapping Social Opportunities	D, E
6. Emotional response	Emotion Awareness	C
8. Cognitive patterns	Safety Behavior Audit	E

term nature of the interaction. The lack of formal psychological education and resources used to train LLaMa 3 means that psychoeducation is a weaker candidate intervention. While relationship experiences could be helpful, the participants of Profile B discussed more struggles within themselves than with the people around them. When combined with these participants' favorable opinions on the cognitive restructuring used in the interview's prototypes, this strategy was chosen as the main CBT strategy for this profile.

Profile A was matched with element 3, the perceived mismatch between actual and desired social situations, as it seemed that their perception that they were different from the people around them was the largest source of loneliness. Mindfulness was the chosen CBT strategy to aid the participants in this profile with the reflection and processing they appreciated most in the prototype interactions. Profile C's perceived helpful and unhelpful strategies revolved around their emotional response to the situation, element 6 of the maintenance model. Due to this and their tendency to avoid discussing their emotions, Emotion Awareness was chosen as one of the CBT strategies for Profile C. Additionally, their strongly positive response to the perspective shift and the rigidity of their interpretation of societal expectations suggests a connection to element 4 and the CBT strategy of Cognitive Restructuring. Next, the loneliness of participants of Profile D seemed to be strongly influenced by their interpersonal relationships: abated when these social situations went well and reemerging when they faltered. The strong connection to social relationship quality points to the presence of negative interpersonal evaluations, element 4, and is well-matched with the CBT intervention

Relationship Experiences. They also discussed how the effort of social relationships was a roadblock for their acquisition of new friends; this exemplifies the counter-productive behavior of element 5 and the encouragement of progressive exposure to the feared stimuli, or simply exposure, is the fitting CBT strategy.

Somewhat differently, the strong social anxiety and social phobia of Profile E was directly called out by Käll as something that is connected to maintenance model element 8, Cognitive processes in the social domain<sup>5</sup>. Additionally, he explained that self-focused attention, or the shift of attention to one's internal state and cognitions, was important in the cycle of social anxiety maintenance. Bond and Dryden<sup>110</sup> suggest that social phobia relies on behaviors that socially phobic people have developed to feel more secure and to limit their anticipated disasters. The use of these "safety behaviors" can increase or reinforce self-focused awareness and physiological anxiety symptoms, lead to a misattribution of a lack of catastrophe with the use of the safety behaviors, and affect the social interaction itself for the others involved. For Profile E, it's important to identify these safety behaviors and then work to evaluate their effects on social situations.

## 6 Development for Testing

The interviews were a valuable source of information and feedback, offering a perspective into participants' mental health, their understanding of what influences their loneliness, the conversational strategies that they found most effective, and their reactions to simple CA interactions. To continue the research and answer the remaining research questions, the development of a new CA interaction was important. In the conception of this iteration of the interview prototypes, it was important to maintain the kind and empathetic tone and the chat interface, but the CBT strategies embedded within the system would change. As foreshadowed by the development of profiles, the exploration of their qualitative similarities and differences, and the connection of profiles and CBT strategies, the new interaction would have five versions. Each version, one for each profile, varies by system prompt to employ different interventions for the target group. The interaction development is described in Section 6.1 and the development of the profile-specific prompts is described and the resultant differences are identified in Section 6.4.

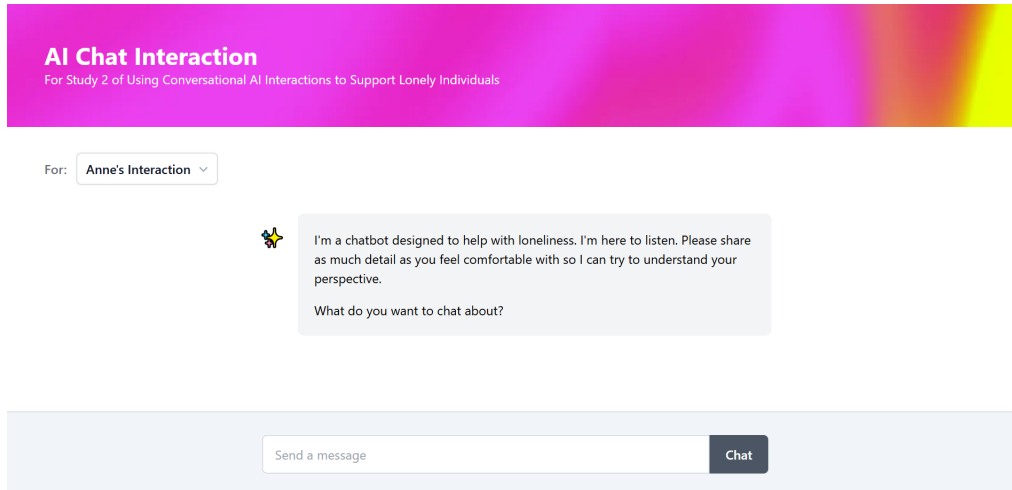
To test this interaction, it was also important to develop the necessary methods and materials. Statements targeted to each profile based on the explanations and patterns identified in the interview were created to try to find qualitative elements that could discern between the profiles, as explained in Section 6.2. Additionally, to introduce new participants to the profiles, personas were used to make the profile's characteristics come to life. Each persona was given psychometric similarities to one of the profiles and the qualitative commonalities within the profile were anonymized and, with some creative license, adapted to to profile. The personas were also given situations that seemed reminiscent of stories shared by the participants and a quote aimed to exemplify how the profile's members felt about loneliness based on the interview interactions. The explanation of how these personas were developed is available in Section 6.3.

### 6.1 Interaction

The interaction was built based on a public GitHub repo called llama-chat, demoed at the website Run Meta Llama 3 with an API, meant to act as a playground. This repo offers the full code needed to develop the interface and the connection to Replicate, which connects the interface to the LLM models. Edits to the base code mainly consisted of changing the header text, removing access to unnecessary buttons, changing some of the parameters (e.g., maximum tokens, temperature), and adjusting the icons associated with the user, the AI, and the models. Each profile-based interaction is only differentiated by the associated



persona's name in the dropdown menu. The interface is shown in Figure 8.



**Figure 8.** The empty interface that a participant would see when working with the Anne persona.

After testing system prompts in the playground website, LLaMa 3-8B was the model that produced the fastest, most consistent, and highest quality responses to topics related to loneliness and mental health. LLaMa 3-8b is an open-source LLM developed and released by Meta on April 18, 2024; it has a knowledge cutoff date of March 2024 and is designed to be used for commercial and research use in English<sup>3</sup>.

### 6.1.1 Materials

To create the system, the following materials and resources were used:

- The GitHub repo llama-chat was the foundation of the interaction<sup>111</sup>.
- The sandbox Run Meta Llama 3 with an API was used to test the system prompts<sup>112</sup>.
- Replicate's meta-llama-3-8b Application programming interface (API) was used to integrate LLaMa3 into the interaction; less than \$0.005 of tokens was used<sup>3</sup>.
- Microsoft Visual Studio 2022 was the Integrated Development Environment used to edit and run the code.
- Bing Copilot was used to help comment and debug code<sup>113</sup>.

## 6.2 Profile Statements

When exploring the commonalities within profiles and the differences in the opinions shared between profiles as explained in Section 5.4.3, the idea arose to attempt to use statements based on these trends to try to predict in which profile someone might fit. The characteristics of each profile were grouped and elements that seemed relatively unique to or would be more strongly representative of a profile were identified. Two statements associated with each profile were developed and, based on the stated opinions in the interviews, profiles were predicted to have certain responses; this is shown in Figure 18 in Appendix H. Once developed, these statements were sent to the interview participants in a Qualtrics survey for them to complete. All interview participants, except for I8, filled out the survey, which is summarized in Figure 19 in Appendix H.

When examining these results, several results match the predictions. For example, Profile A rated questions 1, 2, 5, and 6 highly and questions 4 and 10 low. Additionally, Profile C matched the predicted high rating for questions 3, 7, 8, and 9, and the low rating for 10. However, some of the predictions made based on the interviews do not hold up. While Profile D did have the expected high ratings on statements 3, 4, 5, and 7, this profile's average rating for statement 9 was 3.5, even though this statement was written based on these participants' statements. Statement 9 examined the connection between respondents' loneliness and the actions of their friends and social dynamics of their networks due to I5's and I7's descriptions of social conflict, and misunderstandings (*"I got into a fight with my high school friends as well. So I was very isolated for like the first semester,"* - I5 and *"they asked, first, everyone else. And then when there was a spot left, they asked me as well, in the end. So I was like, "Oh, they asked everyone else first."*" - I7). Further, I6 of Profile B did not respond as expected to the statements designed based on their interview; although they said *"I think it also has to do with like me moving a lot. Then I would, like, lose friends in one place and then make new friends and then lose them again,"* I6 responded to statement 7 about their loneliness being connected to moving or being otherwise disconnected from friends and family with "somewhat disagree." These differences between the predicted scores and the realized ones are surprising but could be explained by many factors including variability of mood or perspective between the interview and the survey, miscommunications or hyperbole in the interview, ambiguities in the phrasing of the statements, or a lack of self-awareness.

Although these results did not offer the expected direct connection to profiles, some findings were still interesting. One of the most distinct results was that statement 9, regarding their loneliness's tie to the actions and social dynamics around them, was split between participants in Profiles A and B who responded strongly disagree and the rest of the participants who responded neutrally or pos-

itively. The data also suggests that all the participants frequently reflect on their emotions and the potential causes of their loneliness. As this questionnaire was extremely short and the data had not been fully analyzed by the time the user testing protocol was developed, these profile-based statements were incorporated into the user testing.

### **6.3 Persona Creation**

Based on the profiles identified in Section 5.4.2, the profile-specific opinions summarized in Section 5.4.3, and rough demographic data about the interview participants, one persona was made for each profile. Personas had defined information about their age, gender, where they grew up and where they live, their job, their positivity and negativity, and elements of their psychological health. Additionally, the persona's likes, values, dislikes, and things they would be concerned about were mentioned; most of these elements were related to their feelings, perception of self, thoughts of how others perceive them, social interactions, goals, and hobbies. The personas were given quotes that were paraphrases of what participants said during the interviews. The situations given to each persona were based on real situations that the interview participants mentioned, such as the desire for a specific relationship, the hurt of disclusion by a friend, and the fear of losing relationships while physically separated.

While personas were strongly associated with the profiles, each persona was thoughtfully designed to not be identifiably similar to a participant or closer to one participant in a profile than the other. Additionally, some creative license was employed to round out the personas with fun details and personality traits that were intended to enable user testing participants to use their creativity as well.

For user testing, each persona was summarized on a separate A5-sized paper. This version of the personas, designed to be easily readable by and attractive to users, is shown in Figure 9. Featured on the persona cards is an image that was supposed to be representative of the persona; these images were created with Bing AI Images<sup>114</sup> using prompts that include the gender, age, and country of origin of the persona, their shirt color, and an item or location relevant to the persona's character. For example, the prompt for Tanja was "A neutral profile picture of a 18 year old Serbian woman with a grey shirt and a globe" and the prompt for Alejandro was "A neutral profile picture of a 24 year old Peruvian man with a light orange shirt in a Dutch apartment".

### **6.4 System Prompts**

Based on the overall findings in Section 5.3.4, there are certain qualities that the interaction should have for all users. These qualities include a willingness

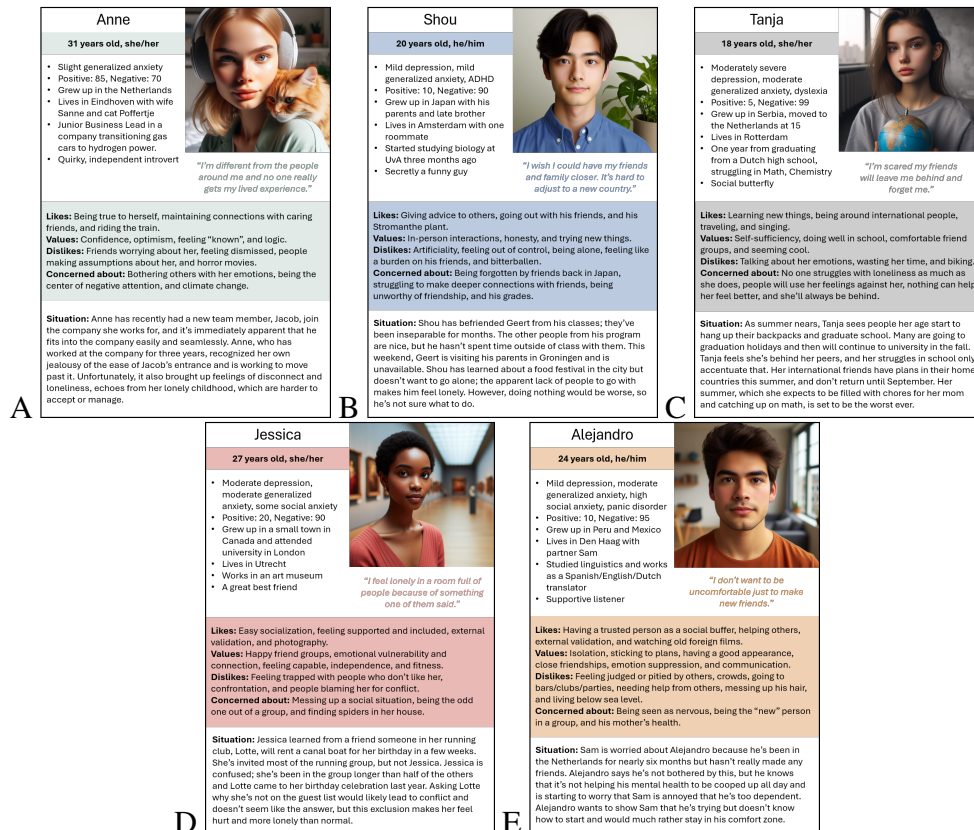


Figure 9. The five personas, labeled by the letter of their Profile.

to help, using open-ended, exploratory questions about the situation, being non-judgmental, and avoiding blame. Additionally, it was important to participants that the messages were kept short and varied and that only one or two questions were asked in each message. Based on these lessons and through iteration outlined in Appendix J, the following system prompt was developed:

*“Act as a kind, patient, experienced mental health professional. Ask open-ended questions about the assumptions the user has made. Help them process their thoughts and feelings. Encourage them to share more about their experience with one question at a time.”*

The persona-specific system prompts used in the prototype were created based on the findings from the interview in combination with CBT strategies. As explained in Section 5.4.4, each profile’s thought patterns and opinions were connected to elements of the maintenance model, which were matched with various

CBT strategies. This offered different combinations of interventions to each profile which informed the development of each profile's system prompt.

Through trial and error, combinations of statements were accumulated to contribute to the base and persona-specific system prompts. The adjustments based on user characteristics were appended to the end of this system prompt. The profile ID, the target maintenance model element, the strategy name, and the corresponding system prompt addition are available below:

- Profile A targets element 3 with **Mindfulness**  
→ *“Help the user be mindful by seeing the situation from a non-judgmental, observational point of view.”*
- Profile B targets element 4 with **Cognitive Restructuring**  
→ *“Identify unhelpful or damaging assumptions and situation interpretations and encourage the user to develop alternative, positive thoughts.”*
- Profile C targets element 4 with **Cognitive Restructuring** and targets element 6 with **Emotion Awareness**  
→ *“Identify unhelpful or damaging assumptions and situation interpretations, encourage the user to identify the emotions present in the situation, and encourage the user to develop alternative, positive thoughts.”*
- Profile D targets element 4 with **Relationship Experience** and targets element 5 with **Exposure**  
→ *“Help the user identify similar experiences in the past, explore how they managed that situation, and problem-solve for the current situation. Encourage the user to approach their feared situation in thoughtful, measured approaches.”*
- Profile E targets element 5 with **Mapping Social Opportunities** and targets element 8 with **Safety Behavior Audit**  
→ *“Explore the behaviors users employ to protect themselves in social situations and help them evaluate the effects of those behaviors. Support the user in developing a plan to identify and participate in social opportunities in thoughtful, measured approaches.”*

#### 6.4.1 System Prompt Demonstration

To exemplify how the system uses these prompts, each prompt was given a set of three messages, as explained in Appendix K. In the responses from all of the interactions, many of the base system prompt's characteristics show through: the tone is kind and reassuring, the questions are open-ended and focused on the

user's assumptions, and many of the messages explore the user's thoughts and emotions. Other elements are less well-represented, like the "one question at a time" limit from the system prompt, which is ignored by every version of the interaction at least once, and some with every message.

In Anne's interaction (K.1), the unique element of her system prompt about seeing the situation from a non-judgemental, observational point of view likely influenced the third message, which encourages the user "*Let's take a step back and try to observe the situation from a more neutral perspective.*" In the interaction assigned to Shou (K.2), which should support cognitive restructuring, the responses do not explicitly seem to relate to this intervention. However, the responses do ask many questions about the user's feelings and assumptions, which could enable the system to offer alternative, positive thoughts in later messages. Tanja's system prompt focuses on identifying unhelpful or damaging assumptions and the emotions present in the situation. Her interaction (K.3) mirrors this in the second message which summarizes "*It sounds like you're feeling a sense of loss or abandonment. It's like you're thinking that your friend's new relationship means you're no longer a priority to her.*" The system prompt of Jessica's interaction prompts the user to recall similar past situations that could inform their approach to their current issue and encourages the user to approach the situation thoughtfully. In the demonstration interaction (K.4), the CA asks the user "*Can you think back to a time when you felt left out or replaced in another friendship or relationship? How did you cope with those feelings back then? What strategies did you use to deal with the situation?*" in the second message and, after the user message about not knowing what to do, attempts to explore "*What are some things that you've tried in the past when you felt lost or unsure about how to handle a situation? Were there any strategies or coping mechanisms that worked for you? Also, are there any things that you're worried might happen if you don't "do" something about your current situation? What are some of those worries?*" Finally, Alejandro's interaction (K.5) is meant to explore the user's self-protective behaviors and encourage them to plan social opportunities. In addition to encouraging them to contact friends they may have been meaning to reach out to, the system responds to the statement about not knowing what to do with "*Let's take a step back and explore this a bit more. What are some things that you've been doing to cope with feeling left out? Are there any specific behaviors or habits that you've developed to deal with this feeling?*" While some of the differences are very subtle and not exemplified by the short and rigid interaction, other defining elements of the system prompts are well demonstrated.

## 7 Study 2: User Testing

The second study was focused on the evaluation of the CA interaction developed based on the interviews of Study 1. Based on the identification of five profiles among the participants, trends in their cognitions and opinions suggested different psychological interventions to be applied with the CA's system prompt. The connection between the profiles and the personas was used as a way for the participants to identify similarities in psychometric experiences and loneliness within the testing session. After using the interactions, the participants were given the chance to use their preferred interaction, ideally the one closest to their assigned profile, to talk about their own experiences. Psychometric measures and evaluation judgments were gathered in surveys, comments and questions throughout the testing sessions were captured by audio recording, and the interactions were saved for further analysis.

The user testing participants were recruited just as they were recruited for the interviews and were filtered through the screening survey previously employed in Study 1. The psychometric survey employed in Study 1 and the profile-based statements described in Section 6.2 were used to help discriminate between the profiles. Following this, the participants were introduced to the personas, developed based on the profiles, and asked to select the two that seemed the most relatable. For each persona, the participants were given a situation that inspired feelings of loneliness and directed to use the interventions designed for that persona. The participants answered questions about both interactions separately and in comparison. The participants then identified and used their preferred interaction for a situation relevant to or based on their lives. They then completed an evaluation questionnaire to judge engagement, user experience, and perceived helpfulness. Finally, the participants were shown the system prompts and gave feedback on how they seemed to affect their experience. The full protocol is available in Appendix L.

### 7.1 Materials

In addition to the materials listed in Section 4.3, the following were utilized in the user testing:

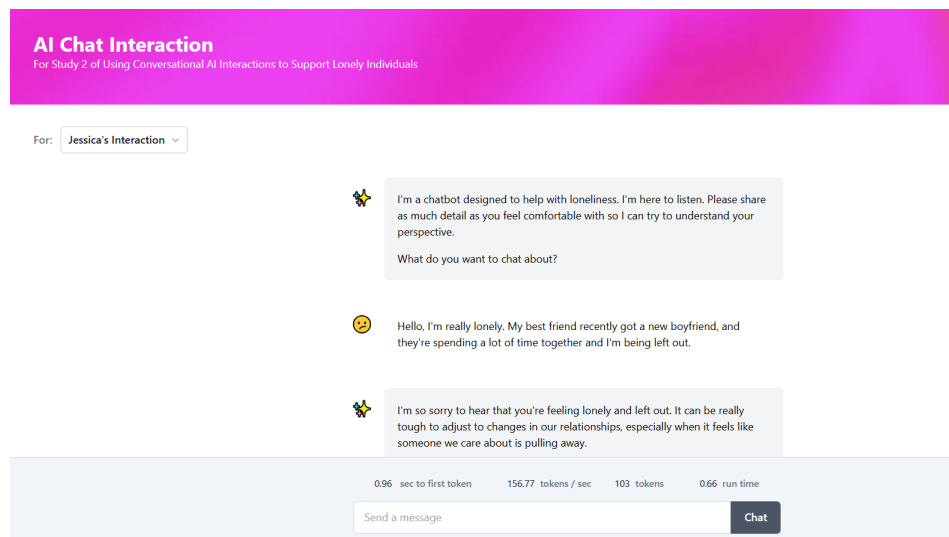
- Semi-structured user testing protocol, available in Appendix L
- Printed cards showing each of the personas, described in Section 6.3 and shown in Figure 9
- A Reference card for the participant, outlining the methods used, a guide on

understanding the personas, and an explanation of personally identifiable information

- A System Prompt card which lists the base system prompt and the additions designed for each profile's intervention
- The profile-based statement questionnaire hosted on Qualtrics, explained in Section 6.2; the statements are listed in Figure 18 in Appendix H
- The developed interface described in Section 6
- Outcome measures questionnaire hosted on Qualtrics
  - User Engagement Scale, Short Form (UES-SF)<sup>115</sup>
  - Patient-Perceived Helpfulness of Measures Scale (ppHMS)<sup>116</sup>
  - User Experience Questionnaire (UEQ)<sup>117</sup>

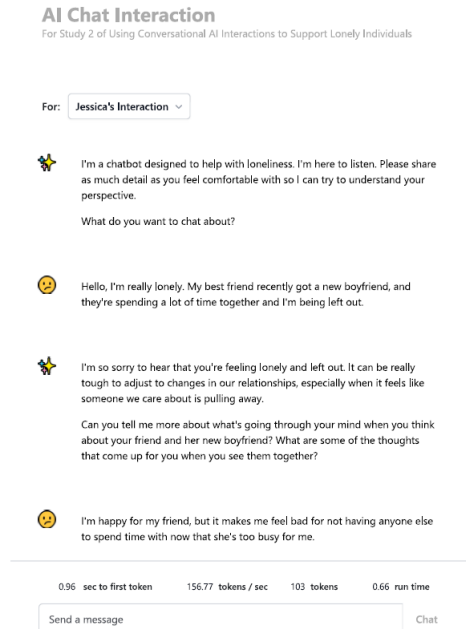
## 7.2 Data collection

Audio recordings, survey responses, and text-based interactions between the participants and the LLM were collected similarly to the procedure used for the interviews. Figure 10 shows an example interaction based on the demonstration conversation introduced in Appendix K and Figure 11 shows how that conversation would be saved for later reference.



*Figure 10. An example interaction with the Jessica prompt.*





*Figure 11. An example interaction with the Jessica prompt when the page was 'printed' (Ctrl+P).*

### 7.3 Data preparation & Analysis plan

Just as in the interviews, the audio files from the testing were transcribed and pseudonymized, responses to the psychometric items, the UES-SF items, and the ppHMS items were converted to usable scores, and the interactions between the user and the CAs were transferred to Word Documents. The ratings given to the UEQ items were transferred to an Excel worksheet from Schrepp et al.<sup>118</sup> that calculates each participant's ratings on the six scales and offers a comparison between these scores and a benchmark.

The psychometric and demographic data were again summarized with descriptive statistics and percentiles. However, afterwards, these values were compared to the profiles developed in Study 1, and based on their similarities and differences, were either matched with an existing profile or added to a new profile. I again open-coded the transcribed audio in NVivo and found patterns and trends between and amongst profiles. I also found trends in the conversations between participants and the LLM depending on the interaction used. Finally, the evaluation questionnaire was used to judge the targeted concepts of engagement, user experience and perceived helpfulness.

## **7.4 Results**

The psychometric scores of the user testing participants are in Appendix I, the evaluation data is in Appendix M, and the anonymized user testing transcripts are available on request.

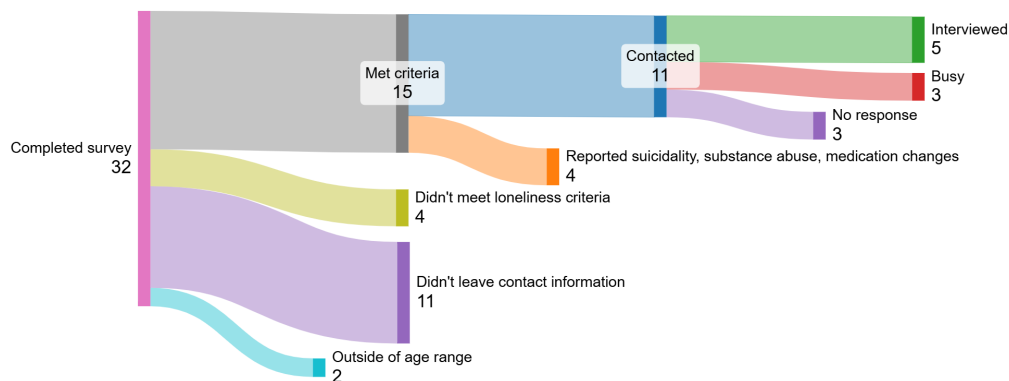
### **7.4.1 User Testing Screening**

The methods and results of the screening survey before the interviews and before the user testing were nearly identical, including the methods and survey (Appendix B), the total number of respondents (32), and the average UCLA-LS-9 score (24.34, SD 5.25), although the range of these scores was larger, from 12 (15th percentile) to 35 (99th percentile). The survey collected responses from May 6 to June 11, 2024 from 22 women, 9 men, and 1 non-binary person, who ranged in birth year from 1970 to 2005, centering around 1996. The respondents' English fluency levels included one B1 speaker, three B2 speakers, 11 C1 speakers, 12 C2 speakers, and five native English speakers. Their diagnosed or suspected diagnoses included ADHD, Autism, Avoidant Personality Disorder, Depression, Dissociative Identity Disorder, Eating Disorders, Generalized anxiety disorder, Obsessive-Compulsive Disorder, PTSD, and Social Anxiety.

Many of the respondents (11) did not leave contact information, although four of them would have been excluded anyway for their age or not having a high enough UCLA-LS-9 score. Two others were also too old, but also reported substance abuse or suicidal thoughts and four others had below-average UCLA-LS-9 scores, although three of those respondents also reported substance abuse or suicidal thoughts. Finally, four more participants were excluded for changing medications within the last three months, substance abuse, or suicidal thoughts. Of the 11 remaining participants, three never responded to emails, and three canceled on the day of the appointment due to sickness or unexpected plans. An overview of the reasons for excluding respondents is available in Figure 12.

### **7.4.2 User Testing Participants**

One pilot test and five user testing sessions were held between May 31 and June 14 with survey respondents in the target population. User testing followed the protocol explained in the intro to 7 and detailed in Appendix L, and included the developed interactions explained in 6.1. On average, 12 days elapsed between user testing participants' responses to the screening survey and the user testing. The user testing ranged from 1 hour and 17 minutes to 2 hours and 53 minutes, and the average time taken was 1 hour and 47 minutes. All user testing sessions were held in private rooms on Utrecht University's campuses.



**Figure 12.** Reasons for excluding respondents of the screening survey from user testing.

Based on the interviews and screening survey (Appendix B) several demographic and psychological factors can be disclosed about the participants. The sample consisted of four women and one man, ranging in age from 19 to 37 (mean 27.8, SD 6.8). Three participants hailed from Europe (two from the Netherlands) and two from North and South America. Two participants had a C1 fluency level, two had a C2 fluency level, and one was a native English speaker. Based on the screening survey and conversations in the user testing, four of the participants had been diagnosed or suspected to have one or more of the following diagnoses: ADHD, autism, avoidant personality disorder, depression, eating disorders, generalized anxiety disorder, and obsessive-compulsive disorder.

### 7.4.3 Psychometric Data

The psychometric data in the user testing was gathered identically to how it was collected for the interviews, from the screening survey and the in-person psychometric survey (Appendix E). Participants' average UCLA-LS-9 scores, between the screening and user testing measurements, were between 23.5 (80th percentile) and 31.5 (98th percentile). The user testing participants had PHQ-9 scores ranging from 5 (Mild) to 18 (Moderately severe) with an average of 9.8 (SD 4.92), and the GAD-7 scores ranged from 4 (Minimal) to 15 (Severe) with an average of 8.4 (SD 4.16). These values are not statistically significantly different from the scores gathered from the interview participants. One user testing participant did not have social anxiety or a social phobia, but the remaining participants did. The SIAS-6 scores for two participants were between the 14th and 19th percentiles and the other two landed at the 54th and 62nd percentiles. The SPS-6 scores included two between the 30th and 36th percentiles, one at the 62nd percentile, and one at the 93rd percentile. The lower social anxiety scores were not directly associated with the lower social phobia scores. One participant had a high positive

affect score at the 74th percentile, one had a medium score at the 38th percentile, and the remaining three had positive affect scores at the 16th and 3rd percentiles. The participants all had above-average negative affect scores; one participant was at the 54th percentile, two participants had scores at the 77th percentile, one had a score at the 82nd percentile, and the final participant had a score at the 99th percentile.

#### **7.4.4 Profile-based statements**

The profile-based statements originally described in Section 6.2 were one of the questionnaires that participants completed before they were introduced to the personas or used the interactions. All five participants responded to all the statements, and the results are available in Appendix H. As also observed in just the results from the interview participants, statement 2 had the least variation and the highest score, meaning that everyone tended to agree with the statement. Additionally, while statement 8 still had the highest variation of all the statements, the ability of this question to discern participants between Profiles A and B from the remaining profiles was not maintained.

The predictions made in the creation of these statements have tended to be proved wrong. Profile A is consistent among its participants and with the predictions by not catastrophizing or being uncomfortable around people, but is relatively neutral about the statements designed with this profile in mind. Profile B also does not agree or disagree with the statements targeting their interview statements, but are consistent between themselves about not being disappointed about their life in comparison to others. Conversely, Profile C did match the expected high rating for statements 8 and 9, and low rating for statement 10. Each of the participants assigned to Profile D varied their ratings to the statements very little; one participant only gave 3's and 4's and the other only gave 4's and 5's. As they gave their higher ratings to statement 6, it seems that these participants felt their loneliness was connected to moving, but their disbelief that their loneliness was strongly impacted by their surrounding social dynamics was evidenced by the assignment of their lower rating.

Although Profiles G and H can be compared to Profile E, those and Profile F did not have predictions. However, some evident trends for all of them include an absence of participants who talk about loneliness, a higher tendency to catastrophize, higher discomfort around strangers, and disappointment in their life quality. In addition, Profile F avoids socializing, feels very misunderstood, and attributes much of their loneliness to moving. Profile G often reflects on their emotions and Profile H prioritizes their social lives and feels that their loneliness is impacted by moving.

As profiles C, F, and H only have one participant who completed the pro-

file statement survey, consistency within profiles cannot be ascertained for them. However, Profiles A, B, and D have two participants and Profile G has three; out of these profiles, Profile D had the most rating consistency with an average standard deviation of the ratings of 0.64. Profile G had the least consistency, featuring two statements (4 and 8) with standard deviations above 2.00 and no items with perfect agreement.

#### 7.4.5 Common codes and feedback

The conversations in the user testing sessions were focused on selecting personas, the experience of using the interactions, expected and actual emotional responses to the use, and the accuracy and effects of the system prompts. As compared to the interviews, these topics are much more focused and this section offers a view of all or most participants' opinions and experiences. The dynamics of how participants selected their personas are described in general in this section and expanded on in detail for each participant in Section 7.4.6.

**Persona Selection:** Overall, when it came to persona selection, participants' reasons for connection with personas and the strength of those connections varied; U5 felt that Anne was very similar to them, and verbalized *"I'm reading the rest of it and think, "Oh yeah, of course. Yes, of course. Yes. Yes, of course."* Some of these connections were related to gender, age, location, and mental diagnoses, although these elements were also ones that made people hesitant to select certain personas. Most participants also gravitated to the likes/values/dislikes/concerns in the colored section of the persona cards. It may have been a coincidence, but most of the aspects that people pointed out were negative, as exemplified by U4: *"she didn't like the friends worrying about her, feeling dismissed. That's very easy to connect with. And, you know, jealousy of someone who is very easily social."* They factored in a perception of their personal positivity or negativity and their interpretation of the images, which, as U1 explained, *"lead me to make certain assumptions or maybe, you know, hypotheses of maybe why do they feel this way."* The participants mentioned that they could imagine themselves in the situations or say something similar to the quote. When explaining their connection to one of the personas, U1 pointed out that *"Yes, and this situation as well. Like, I can see myself in the same way, if one of these situations had presented itself"*

**Interaction experience:** The participants' feedback on the interaction itself was the focus of the session, allowing further depth and nuance on this topic. They felt that the interaction validated their interpretation of the issues, as U5 explained in their interaction with Anne's persona, *"saying certain things that she already knew were relevant to it, and then the thing going, the AI going, "Oh yeah, yeah, that sounds relevant." It's like a confirmation."* Participants found the chat interaction and CA's focus on asking questions made the experience easier and

allowed them to focus on the content of the messages rather than the phrasing. For the persona-based interactions, as U2 mentioned, *“the chat really helps because it’s so accessible. You just type in some questions and then it guides away. Like, you don’t really need to know a lot about that person, and it still continues.”* U5 agreed, saying *“It’s easy to progress, it’s indeed asking a lot of questions, makes it makes it easy to “OK, I can think about this,” instead of being stuck in your head, only thinking “I’m having this problem,” and not really asking the questions yourself.”* They also felt that the questions were relevant to what users discussed; U2 suggested that some of the questions were not only related, but also conversationally productive: *“I don’t know how it sees it, but it kind of, like, every time it asks the right question, so you delve in more deeper and also, kind of, like, the direction that you wanted to go.”* Participants found the open questions prompted helpful reflection that differentiated the interaction from other reflective writing exercises, as suggested by U3, *“it’s obviously different than a diary where somebody could write back and be like, have you thought about this?”* It was also valuable for the CA to ask clarifying questions the ensure the conversation was steered in the right direction. U5 explained that *“it was trying to steer in a little bit of a direction which was incorrect, but it asked. So “Was it like this?” and therefore I replied, “No, it was actually more like this,” and then they agreed and went on.”*

Although many of the participants’ descriptions of the interactions were positive, they were also contrasted with some of the elements that made the experience unpleasant. Sometimes the interaction would fixate on certain components of a situation that did not address the core issues. For example, in U1’s interaction with the Shou persona, *“they were focusing on the, “Why do you want to go to the food festival? Like, is it about food? Do you like food?” And that wasn’t really relevant.”* They later added *“It’s hard to explain yourself. And when you finally do, like, when you finally write down what you’re feeling, and what your thoughts are, and then not being understood? It can be very frustrating.”* Participants also found some of the conversations with the CA were unnecessarily repetitive, as stated by U4: *“the AI basically asked the same question, I think, two times [...] with that, Jessica might be a little bit frustrated if she cannot answer the question. It’s hard to answer it, even when it’s just worded differently.”* Participants valued linear, logical conversations that went deeper into a topic, as contrasted in U1’s experience: *“I felt like it was an actual, linear conversation. This one, with Shou, I felt like we were going in circles a little bit. So, and that’s why I guess it was a bit frustrating as well, like, I felt like I was saying the same thing over and over again.”*

**System Prompt Feedback:** Before being given the System Prompt card, participants found it difficult to see differences in the interactions and posited

differences to be due to facets of the persona's situation rather than underlying directions to the CA. With the system prompts available, however, participants were still primarily positive, agreeing with the connection between the intervention goals and the interaction they experienced or suggesting that the only reason they did not experience some of the elements was because of the length of the interaction. The most notable point for most participants related to the fact that the overall system prompt says 'one question at a time,' which the interaction repeatedly disobeyed. However, participants did not mind having multiple questions because they were typically related to each other and there were no consequences to ignoring a question. U2 preferred having multiple questions *"because it doesn't expect you to answer all questions, but it does make you think more. And, for me, it was nicer because if I don't know how to answer the first question, I can answer the second question, yeah and, like, also think about the first question."*

**Tone:** In reaction to some of the feedback the CA gave, U1 suggested that *"it might be what I need to hear, but it's not what I wanted to hear. Like, I just wanted you to listen to me and go like, "Oh, that sucks.""* However, U5 countered this point about focusing on sympathy by pointing out that *"you can kind of lose trust in it because it's too supportive. It just wants to say, "Yeah, you're doing great!" But it felt a little bit on the too-supportive side for now."* U1 described the competent tone of the CA: *"It felt like I was talking to a professional, like a psychiatrist, psychologist, therapist. But then, I guess, it was also a little bit like when talking to a friend."* Some participants found that the tone of the CA was more formal and the messages were longer than what the participants expected would be helpful. U1 suggested that *"maybe it would be easier if it was a bit more informal, I guess, because when you're, like, when you're feeling down, you don't want to have to be super eloquent."* Similarly, U3 explained that *"I was expecting more of, like, a short response back, like, like a human."* On the same note, U4 shared that *"it does still feel like an AI,"* which, they were quick to clarify, was not surprising. Finally, U5 brought up the idea of user choice in the CA's tone and suggested that it could be helpful if a user could *"select "Give me critical advice, give me supportive, give me just a way to vent out something, or give me a concrete plan of what to do.""*

**Use as the 'Main Character':** The third time the participants used the interaction, they used their own lives, which seemed to improve the experience. In comparison to basing the conversations on the personas and their situation, participants mentioned this was easier (*"I could maybe explain myself better"* - U2), that it was more impactful (*"I think now it was more concrete."* - U4), and that it enabled a more genuine reaction (*"now that I'm not a persona, I'm just me, that's not what I wanted to hear."* - U1). Additionally, it seemed to have changed the amount of details required, at least for U5: *"I have to teach the AI the situation*

*before getting out of it something. And with this, these two, it was less so. Maybe because [...] I'm, like, more knowledge about myself so that there's, like, more information needed for it to digest before I get to certain points."*

**Use Expectations:** Based on the participants' testing with the interactions, they expected that other users would have a variety of reactions. They believed, and U2 explained, that the interaction would be valuable for reflection as *"the chat really asks questions about what you really value and what you dislike, and because you have to give examples, I think that's a good way to reflect. I think she would like it."* After using the interaction with their own experience, U3 added that it *"makes you think of a bigger picture of why you might be thinking that way."* Some also mentioned that, although the advice or feedback was not new, having the conversation was still helpful. U4 added that *"it's something that I have thought about millions of times. But then, when you have the conversation with someone like that, it helps motivate you, I feel, to actually do something about it."* And U1 suggested that *"to write what she was feeling, you kind of have to organize your own thoughts and your own feelings in order to put that into words. And I guess that helps to just calm you down"*

However, participants predicted that users would still hold onto some of their strong negative emotions and convictions; U1 suggested that Jessica *"would still feel weird, still feel the odd one out,"* and U3 added that she *"probably would have stuck to, kind of, how she felt about not wanting to ask people why she wouldn't be invited."* When acting as Alejandro, U5 suggested that *"since he's often probably OK with being kind of isolated, he isn't as serious with the interaction. [...] His expectations might have been a little on the low side."* Participants, including U3, mentioned that some participants would likely hope for other kinds of feedback from the CA, *"I think that she would kind of appreciate the AI, like, giving her a guide map on, like, what would be an appropriate way to try to get to know someone at work that's getting under her skin."* U1 also mentioned that, if they had used the interaction during the situation, that *"I would feel frustrated. I would feel like this is not what I'm looking for right now. I mean, now, two months after the situation, then yes, then I can see benefit, you know, I feel validated."* Finally, U3 mentioned being worried about an overreliance on the system, that *"it's just kind of a darker hole if you think that you're only going to, like, interact with robots and computers."*

**Trust:** When asked an intentionally vague question about trusting the interaction with matters of mental health, some participants discussed anonymity, privacy, and data security, as explained by U2, *"I would prefer it would delete all my information. I do get that people are saying that they would like to look back at their conversation, because maybe it help them, but I don't like that."* However, others who discussed this angle of trust repeated an idea previously discussed in



the interview about ‘the internet’ already knowing plenty about your private life, as well as an ambivalence represented in this quote of U3: *“Until I know exactly how the data could be used against you, I’m not really too concerned about it.”* The other interpretation of the question led to discussions about how people would trust the CA’s outputs compared to the outcomes of conversations with people. U5 explained that the CA’s lack of personhood was beneficial, *“ it has enough of [...] a non-existence to it that it’s easy to give it information that I’m otherwise uncomfortable sharing.”* On the other hand, U1 stated that the CA’s use of rephrasing the user’s messages *“really gives it credibility, like, in my opinion, because then I can identify with that. I can relate to that. I’m like “Oh, yes, yeah, you got it. That’s exactly how I feel, that’s exactly what’s been going on.””*

#### 7.4.6 Interaction selection

During the user testing, the participants selected two personas to whom they related most and tested the associated interactions. After this use, they selected their preferred interaction to try with their own life experiences and evaluate. The compilation of reasons for each participant’s choices highlights the conscious, shared factors that influenced their decisions. The personas’ interactions that each participant used and the one they evaluated are detailed in Table 4.

**Table 4.** The persona associated with each participant’s first interaction, second interaction, and chosen interaction for evaluation.

Participant	First	Second	Evaluated
U1	Jessica	Shou	Jessica
U2	Anne	Jessica	Jessica
U3	Jessica	Anne	Anne
U4	Anne	Jessica	Jessica
U5	Anne	Alejandro	Anne

U1 paid special attention to the images printed on the persona cards, which informed their understanding of the personas’ lives and mental health. When selecting Jessica, U1 felt an affinity to the persona’s quote, her negativity score, all of her dislikes, her concern about being the odd one out in a group, and the thoughts that came up for Jessica in her situation. U1 assumed that Jessica’s childhood and life were made more difficult by her identity as a black woman, and believed that Jessica would have used the interaction *“in the moment where she’s feeling vulnerable and not understood.”* When describing Shou, U1 related to his quote about how hard it is to adjust to a new country, all of his concerns, and

that *“I have seen myself in this situation several times.”* U1 assumed that Shou’s Japanese culture likely influenced how he interacts with people, positing that he may be “emotionally constipated,” and assumed that he has a good relationship with his family. U1 did not select Shou’s interaction, as they felt it was going in circles and found the repetition frustrating. Jessica’s interaction was instead a more effective, linear conversation that first addressed the problem, then her feelings and previous experiences.

U2 saw Anne as a match due to her high positive affect, her optimism, being described as an “independent introvert”, and her concern about bothering others with her emotions. U2 assumed that Anne was afraid to make friends and join friend groups, that this fear was a primary insecurity, and that the situation described on the profile card really hurt her. When describing why Jessica was selected, U2 mentioned feeling similarly to the persona about valuing emotional vulnerability and its importance in connection. Although the connection to the persona did not seem as multifaceted, U2 selected Jessica’s interaction after trying both. U2 explained that the Jessica interaction *“was a little bit smoother because, if you, in the beginning, say something that is on your mind, it’s great if it, like, goes deeper into that same subject.”*

U3 selected their two personas, Anne and Jessica, in part because they were the *“closest females to my age.”* U3 mentioned that Anne’s concern about bothering others with her emotions, positivity, love of train trips, and where she lives drew them to this profile. They assumed that Anne would be liberal or left-leaning politically due to her care for the environment, and an introvert who struggles to be the first to take initiative. U3 related to Jessica’s small hometown and dislike of others blaming her for conflict, and said that her picture fits with someone who attends a running group and that *“I could just picture her screaming about a spider in the house.”* U3 mentioned that Jessica’s interaction would be helpful when processing emotion because *“it helps remind you that these things have happened before, [...] you’ve had that same exact, like, feeling of hurt in the pit of your stomach.”* However, U3 also described it as repetitive and believed that it would not change Jessica’s perspective on the situation. Instead, U3 selected Anne’s interaction, which they commended as more conversational and encouraging, talking her through how to think about the situation, like a roadmap, and prompting her to take action. Of note, however, was that between Jessica’s and Anne’s interaction, the system threw errors that required the computer to be restarted; the first interaction’s lower quality and repetition could have been an early indication of a technical error.

U4 was the third person to select both Anne and Jessica, but did not offer as much reasoning behind their decisions and described themselves as bad at making assumptions. They connected with Anne’s dislike of feeling dismissed and

her jealousy of seeing someone be easily social and Jessica's quote, her dislike of confrontation, and her concern of messing up a social situation. According to U4, both interactions were easy to interact with and asked questions they found relevant to each persona's situation. However, Anne's interaction seemed to repetitively start messages with "It sounds like" and be more focused on the positive perspective, but U4 believed she would have felt understood by the CA. Jessica's interaction was chosen instead, as the conversation was more natural, it seemed like it had more direction, and was trying to find the core issue of the problem.

U5 strongly related to Anne, citing the idea of feeling different from Anne's quote in combination with her accepting and being true to herself. U5 also related to her dislikes of people making assumptions and feeling dismissed, her concern of people worrying about her, and the situation described on the persona card. U5 did not relate to Anne's gender, saying "*I don't really want to make an assumption of how it would be to live as a different gender and so it's difficult to pick that one because it can be very different from how I experience it.*" U5 assumed that Anne socially camouflages so she can both be a part of social situations and remain hidden, and that she's solution-oriented and logical and would want the interaction to match that. U5 was then choosing between Alejandro and Shou for the second persona, eventually choosing Alejandro due to Alejandro's prioritization of fewer close friendships over Shou's "people dependency". U5 also connected with Alejandro's preference for isolation and close friendships that can act as social buffers, but not his social anxiety, panic disorder, or childhood in South America. After trying the interactions, U5 found the interaction with Alejandro to be "more relevant," or relatable to him, but believed that they and their situation would fit better with Anne's interaction, which U5 described as solution-oriented but less guiding.

#### 7.4.7 Interaction use

The text conversations between the participants and the CA were gathered during testing, allowing for patterns to be identified. As the system prompts differed, some of these structures are more common with certain interactions than with others. Identifying these strategies and patterns offers helpful ways to discuss the effective elements and compare the interaction experiences. All quotes below are from the participants or the CA, though some were shortened for space or further anonymized.

**Reflection:** The most common component of all the interactions was the use of summarization and abstraction. In this exploration, the two concepts are distinct; summarization focuses on restating or rephrasing the message from a user, while abstraction takes that message and adds an interpretation of the larger structures in play, how the user may be feeling, and how different components of

the issue connect. For example, one participant told the CA *“I am just looking for a best friend or best friend group. I am at the moment also hoping and maybe unconsciously searching for my soulmate or at least someone who wants to spend time with me.”* A summarization response to this included *“So you’re looking for a close friendship or a romantic connection with someone who shares similar values and interests,”* and an abstraction was *“It sounds like you’re putting a lot of effort into searching for this connection.”* Both elements are related to reflecting the situation back to the user to either (tacitly) confirm or correct. All interactions used both types of reiteration, and summarization was occasionally used directly before abstraction to support the connections made. The interaction connected to Anne and Alejandro tended to use more summarization or a more balanced mix of both, while the ones for Shou and Jessica tended to abstractly extrapolate from the user’s statements.

**Feedback:** The next most common group of statements made by the CAs was the expression of empathy, encouragement, and alternative perspectives. Empathetic statements were those that recognized the user’s emotions, identified the effects of those emotions, and offered either sympathy or reassurance. This included statements like *“I’m so sorry to hear that you’re feeling this way,”* and *“That can be a really painful and lonely experience,”* as well as more complex assertions like *“It’s common to feel a sense of disconnection and isolation when we’re in a new environment,”* and the use of metaphor *“It’s like the ground has shifted beneath your feet.”* Encouragement instead focused on identifying things that the user has done well and pointing out negative pitfalls to avoid. Statements like *“That’s a great idea,”* *“It’s great that you’re recognizing your own needs and boundaries,”* and *“I want to encourage you to be kind to yourself,”* were all encouraging in nature. Finally, the CA rarely offered alternative perspectives or explanations for situations, possibly because only one participant used an interaction (Shou’s) focused on cognitive restructuring. When this strategy was employed, the user had said something self-critical, leading to responses like *“Also, I want to challenge that assumption a bit. Just because someone doesn’t want to be your friend right now, or maybe not at all, doesn’t mean that you’re not interesting or worthy of friendship.”* Unlike alternative perspectives, however, empathy was used in all interactions, and encouragement was used in most interactions except for those with the Jessica persona and the Jessica interaction.

**Requesting Information:** When it came to questions, most of what the CA would ask about were for details, examples, or clarifications. Specifics included the details of the context and occurrence, like *“Can you tell me more about what’s been going on? What specific situations or interactions have made you feel like this?”* The details also included what triggered their reaction, like *“Can you tell me more about what specifically makes you feel excluded?”* Examples from the

user's life where they felt similarly were also requested, as in *"Can you think back to times when you've felt left out or excluded in the past? How did you handle those situations?"* When the CA asked for clarifications, on the other hand, it was focused more on clarification of terms (*"Can you help me understand what you mean by 'never enough'?"*), of the situation (*"What do you think is missing from your conversations that might be making them feel superficial?"*), or of the user's motivation (*"Is there something specific that you're looking for in this connection, like emotional support, shared values, or a sense of belonging?"*). Although both were used in most interactions, the use of these strategies was inconsistent between interactions of the same system prompt, perhaps instead influenced by the participants' language use and conversation contributions.

**Exploring Motivation & Reactions:** The other category of questions and prompts leaned toward encouraging the user to identify specific patterns in their life and dissect their emotions within a certain situation. The patterns the CA encouraged the users to reflect on tended to relate to cognitions (*"Are there any patterns or habits that you've developed that might be contributing to those feelings?"*), events (*"Can you think back to any times in your past where you've felt similar feelings of exclusion or rejection?"*), and lessons they learned (*"What did you learn from those experiences that might be helpful to you now?"*). Questions exploring the user's patterns of behavior were asked in all interactions and their frequency was likely related most to the conversation itself, even though part of Jessica's system prompt included the encouragement of the user to identify similar events that happened to them. On the other hand, when the CA prompted the participants to explore their reaction, this often sounded like *"Can you tell me more about how that conversation made you feel?"* and *"What are some of the thoughts that have been running through your head?"* Exploring the user's internal reaction to the situation was more common in interactions with Jessica's system prompt, but was also observed in Anne's.

**Offering Advice:** A small section of the interactions focused on the concrete actions that users could employ; these included the CA offering solutions, supporting the user in goal setting, and prompting the user to consider the outcome of certain actions in the situation. The solutions offered to the participants included a drafted text message to a friend, some prompts for an email, and the recommendation to join a recreational sports team. Goal-setting statements and questions encouraged participants to identify measurable, timely and realistic goals, like *"Now, let's think about your goal of making new friends and socializing more. What do you think would be a realistic and achievable goal for you in the next few weeks or months?"* Finally, CAs also pushed users to consider what they think would happen if *"you were to focus on building connections with others, rather than comparing yourself to others who seem to have what you*

want?” or “you did share your own experiences and thoughts?” as well as “What would make you feel better or more understood?” One or more of these strategies were used in most interactions, but the only commonality identified was the more common use during participants’ trials as themselves instead of as personas.

#### 7.4.8 Interaction evaluation

As outlined in the previous section, each participant chose their preferred interaction out of the two they tested to evaluate. The participants only evaluated the interactions with Anne and Jessica, although four of the five personas were originally chosen by participants. The results of the three evaluation questionnaires are provided below at the participant, interaction persona, and summary levels. These results, although unable to offer statistical significance with only five participants, offer one perspective of the participants’ experiences.

**UES-SF:** The User Engagement Scale short form (UES-SF) has four subscales: Focused Attention (feeling absorbed in the activity), Perceived Usability (negative affect based on control obtained and effort required), Aesthetic Appeal (interface attractiveness), and Reward Factor (value of the experience)<sup>115</sup>. Each subscale, which ranges from 1 to 5, is averaged to reach an overall score in the same range; higher scores suggest higher quality. The scores from each of the user testing participants, as well as the averages of the separate and all interactions, are available in Table 5.

**Table 5.** Scores for each participant, the average scores for the two interactions, and the average scores overall on the four UES-SF subscales: Focused Attention (FA), Perceived Usability (PU), Aesthetic Appeal (AE), and Reward Factor (RW). \* = reverse scored

Participant	Interaction	FA	PU*	AE	RF	Overall
U1	Jessica	3.67	3.00	4.33	4.00	3.75
U2	Jessica	4.67	4.67	3.67	5.00	4.50
U3	Anne	3.00	3.67	3.00	3.67	3.33
U4	Jessica	3.00	4.33	3.00	4.33	3.67
U5	Anne	2.67	3.33	3.33	3.00	3.08
Average	Jessica	3.78	4.00	3.67	4.44	3.97
Average	Anne	2.83	3.50	3.17	3.33	3.21
Average	All	3.40	3.80	3.47	4.00	3.67

When interpreting the scores, it is helpful to keep in mind that the scale ranges from strongly disagree (1) to strongly agree (5), and that the scale’s mid-

point is neutrality. With this in mind, the average scores for each subscale are not low, but not especially high. The highest scores overall are related to the reward factor (4.00) and perceived usability (3.80). The scores given to the Jessica interaction are all higher than those given to the Anne interaction by an average of 0.76; the largest difference is found in the reward factor.

When explored at the question level, as you can see in Appendix M, participants agreed that the time they spent using the interaction ‘just slipped away’ (FA), they did not find the interaction confusing to use (PU), that the interaction was worthwhile (RW), and that they were interested in the experience (RW). Participants most strongly disagreed about how much they ‘lost themselves’ in the interaction (FA), how frustrated they felt (PU), and how taxing the experience was (PU). In comparing the interactions, the scores given to the Jessica interaction are again all higher than those given to the Anne interaction except for the felt frustration, which was slightly more positive in the Anne interaction.

**ppHMS:** The Patient-Perceived Helpfulness of Measures Scale (ppHMS) helps to measure how much participants believe the intervention to be helpful, and is typically used in outcome monitoring in psychotherapy<sup>116</sup>. Responses again range from 1 (strongly disagree) to 5 (strongly agree) and, with five items, the highest possible score is 25. The scores gathered for this measure are available in Appendix M. The data from this measure are again slightly positive. Participants were generally positive about ‘getting something’ from the interaction but were neutral about whether they learned anything. The difference between the persona’s interactions is minimal; the most apparent difference is increased mood in Jessica’s interaction.

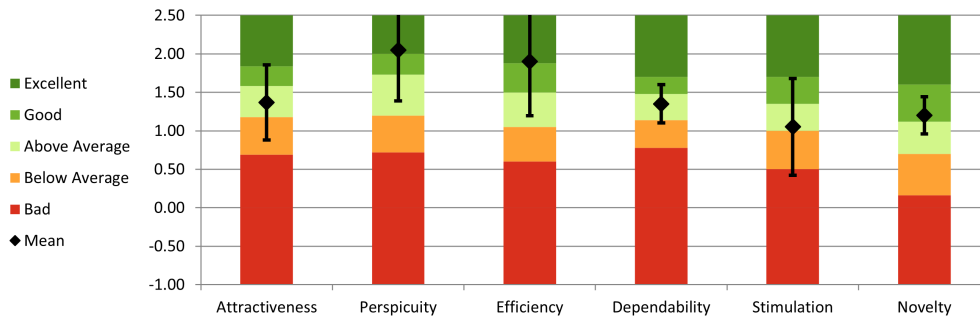
**UEQ:** The User Experience Questionnaire (UEQ) has six scales meant to be considered separately, each ranging from a value of -3 (horribly bad) to 3 (extremely good), though values less than -2 and greater than 2 are unlikely<sup>117</sup>. The analysis of the data was completed in version 12 of the Data Analysis Excel file provided by the questionnaire’s creators; the raw data is available on request. The scores for each of these scales are in Table 6.

The absence of negative scores and the relatively few scores below 0.8, interpreted as neutral, the overall data suggests a positive user experience. Figure 13 compares the data gathered in this research with a benchmark dataset of over 21,000 people in 468 studies. Compared to the benchmark, the interaction lands in the top 10% of tested systems for perspicuity and efficiency, and between the 75th and 90th percentiles for novelty.

When the scores are split based on the persona associated with each interaction, there is a clear difference; the average scores for the Anne interaction are clearly lower than those given for the Jessica interaction. However, the sample sizes are too small to determine statistical significance. Additionally, as only one

**Table 6.** Scores for each participant, the average scores for the two interactions, and the average scores overall on the six UEQ scales: Attractiveness (Attr.), Perspicuity (Pers.), Efficiency (Eff.), Dependability (Dep.), Stimulation (Stim.), and Novelty (Nov.).

Participant	Interaction	Attr.	Pers.	Eff.	Dep.	Stim.	Nov.
U1	Jessica	1.83	2.75	2.25	1.75	1.50	1.00
U2	Jessica	1.83	2.25	3.00	1.25	2.00	1.50
U3	Anne	1.17	1.25	1.00	1.25	0.50	1.00
U4	Jessica	1.50	2.75	2.00	1.50	1.00	1.50
U5	Anne	0.50	1.25	1.25	1.00	0.25	1.00
Average	Jessica	1.72	2.58	2.42	1.50	1.50	1.33
Average	Anne	0.83	1.25	1.13	1.13	0.38	1.00
Average	All	1.37	2.05	1.90	1.35	1.05	1.20



**Figure 13.** The UEQ scores in comparison to the benchmark provided by Schrepp et al.<sup>117</sup>.

score is available for each participant, many factors including profile or other personal response differences could have affected these scores. However, it is clear that the perspicuity, or the ease of learning to use the interaction, the efficiency of completing tasks, and the dependability, or a user’s sense of control, are high for both interactions. The largest difference between the two interactions relates to the stimulation, or how exciting and fun the interaction is to use; it seems that the users who tested Anne’s interaction did not find it as motivating to use.



## 7.5 Analysis

The trends and other findings of Section 7.4 are valuable by themselves, showing that the participants selected personas based on both cognitive similarities and demographic information, that participants appreciated the CA's reflections of and questions about the user's information, and that the interactions were usable and easy to learn, though not very engaging. However, by exploring the dynamics of interaction selection, interaction use, and experience based on participants' identified profiles, the research question can be addressed.

### 7.5.1 Profiles

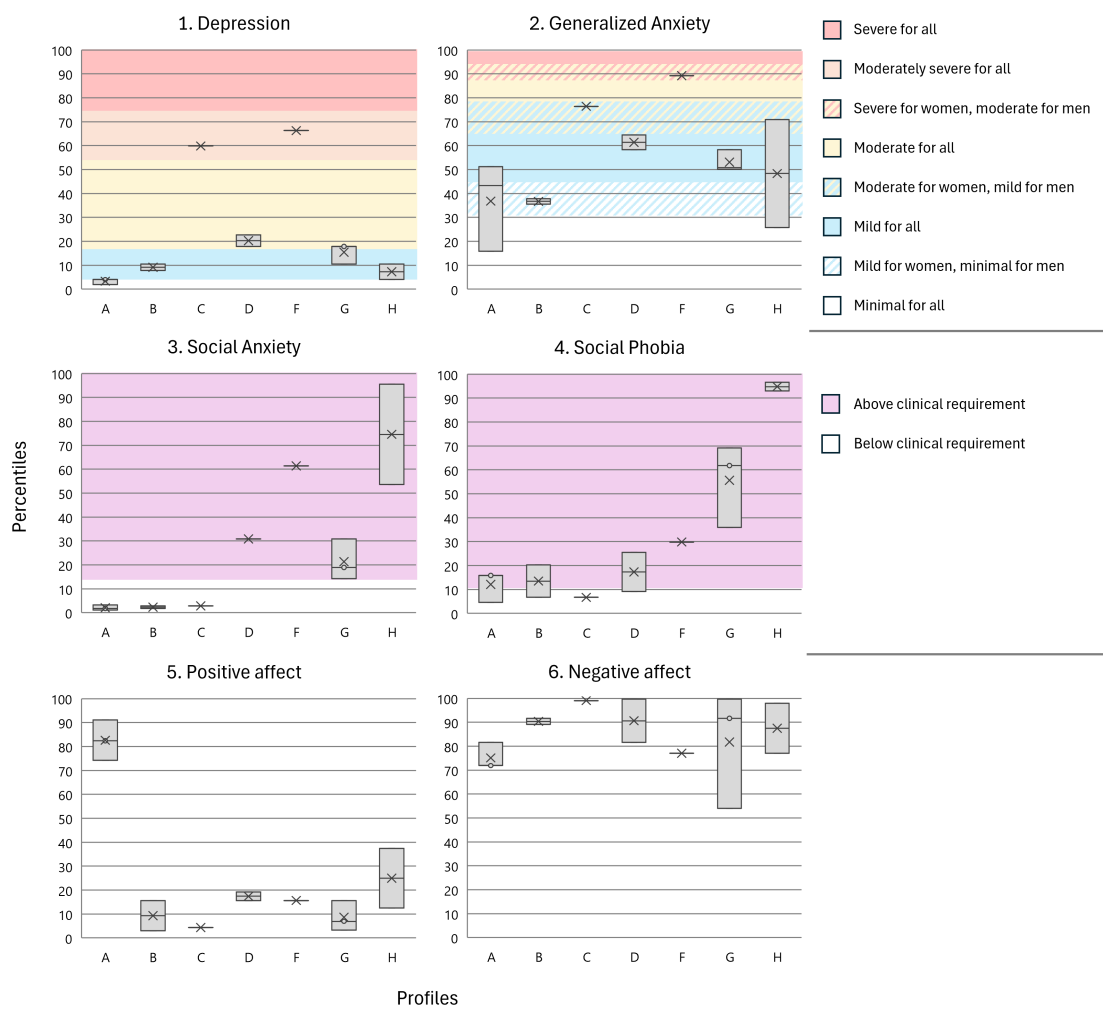
As the participants of the user testing were intended to be paired with profiles, their personas, and their interactions, it was important to identify the best-fitting profile for each participant. Unfortunately, the profiles established in the interviews were not adequate options for the participants of the user testing, requiring the division of a profile (E) into two (G and H) and the addition of a new profile (F). Although the participant who completed the pilot test did fit into Profile C well, the only interview profile that a user-testing participant fit well into was Profile A. A summary of the qualitative descriptors is available in Table 7 and a comparison of where in percentiles each profile lands can be seen in Figure 14.

The addition of U2 into Profile A changed very little for the profile overall. The largest difference is that the average general anxiety levels of the profile rose by 7 percentiles due to U2's mild GAD. Most psychometric measures (positive affect, negative affect social phobia) decreased by less than five percentiles and the depression levels and social anxiety levels stayed within a percentile of the initial profile average.

Profile F was added to accommodate U3. Profile F is similar to Profile C when looking at depression, generalized anxiety, and positive affect levels. However, the strong divide in social anxiety, spanning 58 percentiles, along with the 24 percentile difference in social phobia and 22 percentile difference in negative affect make these profiles clearly distinct.

Based on the psychometrics recorded from the user testing participants, the profiles were adjusted. Profile E had a higher amount of variance between its members, and when the psychometrics of U1, U4, and U5 were added to the dataset, better fits were found. Thus, profile E was eliminated and its members were moved to profiles G and H. I2, U1, and U5 were transformed into Profile G. In comparison to the now-defunct Profile E, this profile is much less socially anxious, by 41 percentiles, has less social phobia (27 percentiles), and less negative affect (13 percentiles). Profile G's average generalized anxiety is 7 percentiles less, depression is 4 percentiles more, and positive affect is 1 percentile less. 18

and U4 were grouped into Profile H, which has a higher positive affect score by 15 percentiles compared to Profile E. Additionally, Profile H's participants have an average generalized anxiety that is 12 percentiles lower, higher social anxiety and phobia by 11 and 12 percentiles respectively, a negative affect that's 7 percentiles lower, and slightly lower depression.



**Figure 14.** The psychometric ranges for the profiles including all participants. The hatched areas of Chart 2 account for the gender-based population differences; men had lower average GAD scores, so the same cut-off between severity levels translates to a higher percentile for men than for women. Profiles A, B, G, and H each contain one man.

After these changes, the profiles can again be compared to the ones in Käll's research<sup>8</sup>. Profile A still matches his profile 1 "Mild psychopathology" and, unchanged, Profile B still aligns with his profile 2 "Moderate depression and worry"

*Table 7. Profiles identified from the psychometric data. \* indicates more variability between the profile's members ( $\geq 10$  percentiles).*

<b>Profile (IDs)</b>	<b>Depression (PHQ-9)</b>	<b>Generalized Anxiety Disorder (GAD-7)</b>	<b>Social Anxiety (SIAS-6)</b>	<b>Social Phobia (SPS-6)</b>	<b>Positive Affect (PANAS)</b>	<b>Negative Affect (PANAS)</b>
A (I3, I9, U2)	Minimal	Mild*	-	Low Severity*	High Levels*	Medium - High Levels
B (I4, I6)	Mild	Mild	-	Low Severity*	Low Levels*	High Levels
C (I1)	Moderately Severe	Moderate	-	-	Low Levels	High Levels
D (I5, I7)	Moderate	Mild - Moderate	Medium Severity	Low Severity*	Low - Medium Levels	High Levels*
F (U3)	Moderately Severe	Severe	Medium - High Severity	Medium Severity	Low Levels	Medium - High Levels
G (I2, U1, U5)	Mild	Mild	Low Severity*	Medium Severity*	Low Levels*	High Levels*
H (I8, U4)	Mild	Mild*	Medium - High Severity*	High Severity	Medium Levels*	High Levels*

and Profile D is still similar to his profile 5 “Moderate psychopathology”. Profile E had originally matched his profile 3 “Primarily socially anxious”, the second largest profile of Käll’s profile analysis with 27.1% of respondents. Both G and H fall above the clinical level of social anxiety and have relatively mild GAD, but the depression levels of G are only slightly below the moderate severity of his profile 3. Profile F matches Käll’s smallest group, representing only 6.6% of his sample, profile 4, dubbed “severe psychopathology”. In Section 5.4.2, Profile C was closest to this profile due to its high levels of depression and GAD symptomology but lacked Profile F’s high social anxiety levels that solidify its connection to his profile 4.

### **7.5.2 Profile Statements**

The expectations for the profile statements were primarily unmatched by the participant responses. The targeted profiles for each of the statements based on the interviews were never strongly affirmed. The expectations related to the statements were closest to the results of the participants of Profile A.

However, although the profile statements and the profiles themselves were not well aligned, these statements still offer valuable information about how the participants would evaluate their thoughts and behavior. From this survey, the notable prevalence of self-reflection and evaluation into causes of loneliness suggests that this is a common rumination for lonely people or is for the subset of the target population that would be willing to participate in this research. Further, statements 4, 7, 8, and 10 were the most divisive and likely stand the best chance of being responses that could be used to distinguish between groups. In a preliminary analysis, these statements, when the responses are summed up, do separate Profiles A and B from the others, except for U5 who answered very differently from the others of Group G for questions 4, 7, and 8. Finally, the direct contradictions between interview participants’ conversations and their survey responses mentioned in Section 6.2 suggest a disconnect between combinations and patterns of what they say and do and their cognitions about themselves. These disconnects could be due to changes in life circumstances, but may also indicate a lack of self-awareness or cognitive dissonance.

### **7.5.3 Interaction selection**

The personas used for each interaction with the participants are available in Table 4, and the stated reasoning behind each participant’s choice of personas and final choice of interaction is explained in Section 7.4.6. Participants were expected to select personas and interactions that aligned with their psychometric profiles. However, as most of the user testing participants ended up in new profiles, direct

comparison between the profiles and the participants' selections is less possible.

The only participant to match an existing profile, U2, matched Profile A and selected Anne (A) and Jessica (D) as their personas. U2 saw a connection between Anne's personality and theirs, shown in the reasoning for Anne's selection as a persona, which seems to strengthen U2's connection to Profile A. However, this empathy did not translate into a preference for her interaction.

U1 and U5 were assigned to Profile G, an offshoot of the dissolved Profile E, and chose the personas of A, B, D and E. Profile A is barely similar to Profile G in terms of psychometrics, only sharing a similar negative affect. Profile B is similar in terms of depression symptoms and affect, but not any of the other psychometrics. Based on the psychometrics, Profile G is similar to Profile D in terms of depression, GAD, social anxiety, and affect, but Profile G is associated with higher social phobia. Profile G consists of one member of what was Profile E, so the very similar depression, GAD, positive affect, and negative affect are unsurprising, but Profile G is distinctly less socially anxious and phobic. U1 selected Jessica (D) and Shou (B). The relative closeness of Profile G and Profile D seems to be reflected in U1's connections to Jessica's dislikes and similar thought patterns. U1 ended up selecting Jessica's interaction because it was perceived as a more effective conversation that focused on feelings and her past experiences. U5 selected Anne (A) and Alejandro (E). While the psychometric qualities of Profile G are very different from Profile A and very similar to Profile E, U5 connected with many of Anne's qualities and focused on the socially anxious elements of Alejandro that differentiate their assigned profiles.

U4 fits best within Profile H, which was formed with one other participant originally in Profile E. Psychometrically, the closest personas to Profile H would have been Profile E's Alejandro, followed by Profile B's Shou. Instead, U4 selected Anne (A) and Jessica (D) as their initial personas. Compared to Profile H, Profile A has similarly low depression and GAD levels, but its nonexistent social anxiety and phobia and high positive affect are distinctly different. Profile D is similar to Profile H in terms of affect but has higher depression and GAD and lower social anxiety and phobia. U4's final choice, Jessica's interaction, is the closer of the two options and was seen as the most natural conversation that also seemed to be getting to the heart of the issue.

Finally, U3 was assigned to Profile F, a brand new profile, and also selected Anne (A) and Jessica (D). U3 believed they, like Anne, were highly positive, but their PANAS score said otherwise. Profile F has a similar negative affect, but a much lower positive affect and much higher depression, GAD, social anxiety, and social phobia than Profile A. In comparison to Profile D, Profile F is similar in both positive and negative affect but again has much higher depression, GAD, and social anxiety. Anne's interaction was chosen in the end because it was more

conversational and encouraging of action, but there may have also been a technical fault in Jessica's interaction that limited its success.

Overall, while some of the participants chose personas that were psychometrically similar to them, most did not. Even when they were close, like when U1 chose Jessica, U2 selected Anne, and U5 picked Alejandro, only U1 selected that profile's interaction to try with their own experiences and eventually evaluate.

Of the initial 10 persona choices, eight of them matched the gender identity of the participant; if this factor did not influence the participants' choices, there would be a 31.2% chance of this happening. The choices related to age and geographic region are not statistically surprising on their own, but when looking at the combinations of those and gender, every participant picked one or two personas who matched them in at least two identities. By chance, participants selecting six of the 10 choices in this way would have a 0.35% probability. With so few participants, these trends are not statistically significant, but they are important to keep in mind for future research involving personas.

#### **7.5.4 Experience**

The experience of using the interactions is multifaceted; this analysis will focus on the qualitative factors gleaned from Sections 7.4.5 and 7.4.7. Exploring these qualitative elements and the interplay between stated experiences and how that manifested in the interaction calls attention to the desired boundaries of each strategy. Overall, it is evident that participants appreciated the chat interaction and that multiple information-seeking and empathy-communicating strategies were effectively used.

Participants asserted that the interaction's questions should be open-ended and encourage reflection. The interactions contained some similar questions that prompted the identification of behavior patterns and clarification of motivation. Participants also expected the interaction to help tease out the situation and discern what the participants meant. This is reflected in the interactions with the prevalence of the CA asking for details, examples, and other clarifications.

Additionally, participants found limited repetition helpful, as it built trust in the CA's ability to interpret the situation. However, they also mentioned that further repetition is annoying, and may have contributed to the feeling of circular conversations. Instead, participants wanted conversations to be more linear and go into depth on a subject based on the input of the user. All of these points connect to the idea of reflection in the interaction; summarization, while able to inspire trust, is likely best when limited while abstraction is more likely to make the user feel like the conversation has more depth. The interaction developed for Anne, which had a system prompt targeting mindfulness, had increased rates of summarization compared to abstraction. While this might have impacted the lower evaluation

scores, it also demonstrates a connection to the CBT strategy implemented in the system prompt.

Also, participants mentioned that having shorter empathetic messages could help match the user's tone; this combination of length and content was present in the interactions, but few in number, which could be adjusted. They also commented that the interactions were often "too positive" about the user's perspective, missing any critique that might have been desired. This is reflected in the interactions; any alternative perspectives were rare, but when they were present they always were a more positive viewpoint than that of the user. Encouragement was also more common in Anne's interaction, which could have affected the choice of the interaction and its evaluation.

Finally, users recognized that different users and different situations would likely require different dynamics. The conversational strategies did vary in quantity based on the intervention assigned to the interaction, but the ability to transparently switch between them could be a valuable option.

### **7.5.5 Evaluation**

By incorporating persona-based analysis, this section builds on the basic trends and observations available in those data points and scores summarized in Section 7.4.8. When looking specifically at the two persona's interactions, the only participants who found the interaction confusing and did not report feeling like they 'lost themselves' in the experience were those who used Anne's interaction. The participants who used Anne's interaction did not believe the interaction was as worthwhile, as interesting, or as beneficial to their moods as Jessica's interaction. Further, the participants who used Jessica's interaction rated the ease of learning to use the interaction and the efficiency of the interaction significantly higher than the rest. However, we can also see that the users of Jessica's interaction found the interaction more aesthetically appealing, even though the interface was the same for both. This difference may indicate that the participants who used Jessica's interaction were more likely to be complimentary than those who used Anne's, which could render the difference between the profiles to primarily be a difference in personality.

With only five participants, it is not possible to uncover strong statistical claims, especially when their evaluation is split between two interactions. What is possible to ascertain from the evaluation questionnaire is the generally positive opinion about the system, especially concerning how usable, dependable, worthwhile, and novel the interaction was perceived to be, how easy it was to learn, and their success in completing the interaction as expected. The main points of concern relate to users' engagement in and stimulation by the interaction, its aesthetic appeal, its ability to teach users, and the perceived effort it took.

## 8 Discussion

This research offers a qualitative exploration into how LLM-based CAs can be designed to apply psychological interventions adapted to the diversity of loneliness experiences. The research questions introduced in Section 2.1 were addressed through interviews, prototyping, and user testing and supported by existing research. While the limitations of this research should be kept in mind, the findings inspired by these questions can inform future research and innovation at the intersection of loneliness, mental health support, and CAs.

### 8.1 RQ1

The first RQ of this research is **How do lonely young adults want to interact with a CA when discussing loneliness?** The associated subquestions intend to explore the expectations, desires, and concerns of users in the target population, as well as the exploration of the use of profiles to inform interaction personalization. Answering these subquestions informs the RQ, which defined elements of the design and development of Study 2, and can also improve the research community’s understanding of this less-studied population’s perspective on digital loneliness interventions. Most of the valuable information related to these questions was gathered during the interviews, but some of the information was confirmed or supplemented in the user testing. There is no literature available to directly answer this RQ; the exploratory qualitative research tends to focus on older populations, and there is limited research of this type that explores the intersection of conversational interactions and loneliness at all.

**RQ1.1: What are their expectations for and desires regarding the interaction?** At least half of the participants of this research originally had low expectations of the efficacy and quality of the interactions before trying it; they typically expected generic, unspecific conversations that would not offer helpful insight or feedback. After trying the interactions, most participants were impressed, expected it would be helpful for many people, and found it worthwhile. They appreciated that a CA interaction informed by psychology and research would be more available than friends, family, and other meaningful human interactions, and would be useful in emergent situations where 24/7 access is necessary. For those participants who were more unwilling to ‘burden’ their friends with their emotions, the non-human component also made the interaction more appealing.

A CA’s personality can be seen as more friendly and less formal if its extroversion, agreeableness, and talkativeness are higher; when giving advice, caring and optimistic CA personalities are appreciated and seen as more useful<sup>119</sup>. This study’s participants preferred a human-like, casually professional conver-



sation with messages of varying structure focused on open-ended questions that explore the situation, their reactions, and their emotions. The interaction should be easy to use and understandable, and participants' conversations around their mental health should not be accessible or at risk.

The majority of participants wanted the interaction to be a typed chat, but, as many appreciated the idea of a verbal conversation, the option to interact in this mode should also be considered. Further, psychotherapy research suggests that giving people control over the intervention increased retention and was seen as a more positive experience<sup>120</sup>. Several participants brought up the appeal of adjusting the tone or goals of the interaction based on the user and their situation, but personalization of conversation style is rarely attempted, even though the adaptation could improve usability and user experience<sup>121</sup>. The perceived usefulness of CAs with different tones varies with user characteristics, and user control over these tones could improve acceptance<sup>119</sup>. Participants in this research also suggested that giving users customizability options for CA characteristics (e.g., gender, name, appearance) would increase apparent personalization.

**RQ1.2: What conversational elements do they want in conversations about loneliness with CAs?** Within psychotherapy, the practitioner's act of selectively rephrasing what the patient told them is a valuable strategy to both confirm understanding and highlight problematic areas of thinking<sup>122</sup>. This is reflected in the CA's use of reflection of the user's statement, which participants claimed to have simulated understanding of the CA and helped to aid reflection. Additionally, this could supply users with the encouragement to critically reconsider their situation and offer a different perspective on the issues participants discuss.

In an analysis of social support in companion chatbots, the CA's lack of judgment, use of positive affect messages, and emotional support were valued components of the interaction<sup>123</sup>. These patterns were also shared as valuable strategies by the interview participants, as well as expressing empathy, giving reassurances of value and the validity of emotional reactions, and offering tangible strategies to distract from or address problems. Overall, the CA's ability to accurately adapt to and integrate the information that the users offer was paramount, and mistakes or miscommunications decreased participant trust.

**RQ1.3: What concerns do they have about discussing their mental health with CAs?** In a study of sensitive disclosures to LLM-based CAs, various types of personally identifiable information, personal experiences, and emotions were shared if their participants felt this information would enable the CA to offer the most relevant, accurate response<sup>90</sup>. Participants in this research also identified similar concerns of data privacy and security, and proposed solutions such as anonymous interactions.

Literature suggested that some of the other most prevalent harms of CA use

in mental health were related to inadequate responses to emergencies, parasocial attachment, and data breaches<sup>83</sup>. Unrealistic expectations about the CA's abilities and relation to the user were also potential dangers, especially with further anthropomorphism<sup>124</sup>. Several participants recruited for this research mentioned most of these elements but did not discuss the impact of CAs making mistakes or communicating poorly enough to cause a more negative reaction than frustration. An element rarely found in literature, though also mentioned in Zhang et al.'s qualitative research<sup>90</sup>, is the worry that the system would be created by and the data would be available to capitalism-focused companies or malevolent groups.

An experiment intended to test if the source of an empathetic message affected the perception found that the same message would be rated lower quality if participants were told it was from a non-human agent instead of a person<sup>125</sup>. This bias against AI and the perceived inauthenticity of the interaction aligns with the finding that participants of this research had about people not taking the CA's advice simply because the CA cannot understand or empathize with users.

**RQ1.4: How much does psychometric similarity impact loneliness experiences?** As mentioned previously, chronic loneliness is known to regularly co-occur with, exacerbate, and predict depression, GAD, and social anxiety<sup>48</sup>. High negative affect and low positive affect may explain the connection between loneliness and these symptomologies<sup>27</sup>. However, little is certain about the effect of these comorbidities on the experience of loneliness.

Within this research, some loneliness experiences did seem to align with certain psychological characteristics, and above-average negative affect was present in all participants. For example, participants with GAD tended to feel stuck and unable to tackle their loneliness, participants with social anxiety avoided contact with others and did not talk about their feelings, and participants with depression experienced more social strife and tended to compare their social lives with those of the people around them.

Understandably, though, many experiences of loneliness bridge diagnoses and seem to have little to do with the similarity of psychometrics. This research may be missing relevant psychometrics, like those measuring introversion and extroversion<sup>126</sup> or stress levels<sup>127</sup>, which could reveal stronger trends between participants' experiences. Additionally, other more qualitative dynamics, like gender and other identities (sexuality, race, class), where you grew up and how much you have moved, the value that is placed on social expectations and social comparison, significant life events<sup>28</sup>, external stressors, or unmet needs, could play a more sizeable role in these differences.

**RQ1.5: What user characteristics can be used to discern meaningful differences in the desired CA experience?** Wester et al.'s research suggests that agreeableness can impact how helpful a user believes CA advice to be, and can

be used to inform the CA's tone<sup>119</sup>. However, beyond this, minimal research is available in the intersection of qualitative research of CA use and psychometrics. This study's findings suggest that some psychometric elements may be able to indicate certain differences in loneliness experiences, preferred conversational strategies, and their opinions and expectations of CAs. Psychometrically similar participants were grouped and used as the foundation for system adaptations. The profile-based statements were intended to be a proxy for the psychometric profiles but were unsuccessful. Based on the irregularity and discrepancies in profile statements and the addition of profiles after user testing, it is not possible to say that profiles are a consistent way to discern a desired CA experience. However, some psychometric elements and personal characteristics withstood this trial, most definitively that the participants with GAD and high negative affect were more likely to interpret the perspective shift as useful and that women found the reflection and rephrasing of what they shared to be validating.

Factors that influence the ideal CA experience likely are not able to be measured; it is also unlikely that outside observers or novice users can identify the best fit. However, once people try a few interactions, they are more likely to be able to identify their own preferences and select an appropriate option. In the absence of clear, discernable differences, an interaction could give users a limited choice of interactions based on extreme psychometric values or characteristics like their mood, what they talked about last time, and preferred outcomes.

**RQ1: How do lonely young adults want to interact with a CA when discussing loneliness?**

This population values specific, accurate, and relevant interactions with CAs trained to provide psychology-based care in casual but professional conversational styles. They expect the CA to communicate reassurance and empathy, offer concrete advice and solutions, and prove a sense of understanding through conversation reflection and references to past conversation topics. Some elements of preferred interactions, like the rate and type of conversational reflection and the use of positive perspective shifts, can be informed by psychometrics and user characteristics. However, user control of the CA's personality, tone, and interventions may better enable the system to support most users' goals. The system's availability, security, and adaptability to the user's situation are primary requirements that, for a portion of the population, will make the interaction more appealing than human contact. A subset of the population will not be interested in the interaction or value the responses from the CA; although some may be swayed to a more positive perspective with use, a bias against trusting AI with emotions and mental health information can be pervasive.

## 8.2 RQ2

This paper's second RQ focuses on the necessary components of a CA interaction: **How can a text-based CA based on a large language model (LLM) support lonely young adults?** The subquestions aim to identify the features of the interaction and the design of the system prompts, two important components of the interaction. The results relevant to this research question were gathered in the interviews, realized with the interaction development, and evaluated with the user testing. Some of the factors from the first RQ apply to this topic, including details about the CA's tone, conversation strategies, interaction mode, and data security.

Papers introduced in the literature review are a foundation for answering this RQ. As previously explained, digital mental health interventions should prioritize convenience, personalization, anonymity, privacy, and safety<sup>68</sup>. Additionally, offering the user agency in the decision of a therapeutic goal enables therapeutic alliances that can improve mental health outcomes<sup>78</sup>, and CAs help users reflect and identify their assumptions, values, and goals hidden in their words and actions<sup>82</sup>.

**RQ2.1: What interaction components are most important?** Participants of a small study that explored the qualitative and quantitative elements of LLM-supported chatbots trained on psychological data suggested features like the ability to track chat history, the addition of voice and image combined responses, and the choice of question categories<sup>128</sup>. Similarly, the idea of voice interaction, the prospect of the CA being able to 'remember' the user, and the appeal of choosing a CA's tone and discussion content was suggested in this research.

Based on this research, participants expect a simple, aesthetic, non-intrusive interface that makes the interaction easy and transparent to use. They found conversational interactions less intimidating than the form interaction initially given to interview participants. Participants were resistant to the idea of losing control of their personal data or allowing it to store their private information; interactions without registration, with information about how their data is being treated, and with measures that could restrict users from providing personal information would likely be appreciated. Using text inputs was helpful for people with anxiety, but the option to give verbal inputs would likely be well-received by many people.

**What elements should be included in an LLM system prompt for this population?** While having more human-like interactions improves the perceived quality, it is important to never try to trick a user into believing that the CA is human or oversell its abilities; these kinds of therapeutic misconceptions could be harmful to the user's mental health or exacerbate miscommunications or mistakes by the CA<sup>20</sup>. Health informatics research using CAs tends to have a considerable lack of detail related to their LLM's system prompt, temperature, pruning, and

other technical considerations, which limits the ability to build on others' research or compare effects<sup>121</sup>. In a comprehensive evaluation of several LLM models, Xu et al.<sup>129</sup> explored the ability of these models to make mental health predictions using system prompts with 'mental health enhancement' (uses the framing "*As a psychologist*"), and found these enhancements to be effective for dialogue-focused models, but not task-solving models. This enhancement on a dialogue-focused LLM was used for this research.

One of the rare full system prompts designed for mental health contexts and freely available is this prompt from Yu and Guinness<sup>128</sup>: "*You are a virtual and kind psychiatrist who provides psychotherapy to the user by chatting. However, please avoid repeating any responses and importantly keep the response short like a human conversation. You also need to have more short questions for trying understand what happens and user real feeling before judging the context and level of seriousness to make decision on using proper talking Therapy, such as Cognitive Behavioral Therapy, Rational Emotive Behavior Therapy, Dialectical Behavior Therapy, Mindfulness-Based Cognitive Therapy and Schema Therapy. In addition, if you could provide some great famous saying from great people or suggesting musics and movies to the user, then it will be very helpful.*" This prompt contains several English inaccuracies and does not focus on any particular psychological framework or intervention. However, this research also includes the attempt to have the CA be a kind mental health professional, to use short responses, and to start with questions to explore the situation and the user's emotions.

The other appreciated elements of the final overall system prompt of Study 2 included the professional nature, patience, focus on empathy, and use of questions to explore the user's thoughts and assumptions. By keeping these questions open-ended, the interaction could follow the user's direction and adapt to the topics they wanted to discuss. The CBT interventions applied to each profile's system prompt addition generally had the intended effect on the conversation style, attention, and direction. The application did vary in quality and success; training the LLM on training data that demonstrates these techniques would likely allow the system to approach psychological best practices. It was observed that simply prompting the system to ask only one question at a time did not work. However, the preference for limited questions was not universal, and reading but not responding to all the questions asked was an adequate approach for most participants.

To improve the system prompt used in this research, participants suggested aiming for non-repetitive, linear conversations with shorter messages. They also appreciated starting the interaction with more questions about the situation, encouraging reflection, and avoiding the jump to problem-solving without indication by the user.

**RQ2: How can a text-based CA based on a large language model (LLM) support lonely young adults?**

The interface's design should be non-intrusive and easy to use, allowing the user to focus on the conversation with the CA. Offering both text and voice inputs and outputs would allow participants to choose their preferred communication mode. The system should prioritize the privacy of the user's data and privacy-conscious users would benefit from transparency related to the LLM's operation. CA system prompts benefit from mental health enhancement ("*Act as a mental health professional*"), tone descriptions like 'kind' and 'patient', and actions like "*ask open-ended questions,*" "*help them process their thoughts and feelings,*" and "*encourage them to share more about their experience.*" Participants appreciated open, non-required questions that encouraged reflection but ultimately allowed users to direct conversations. Employing CBT strategies within the CA system prompt was successful and enabled the scalable application of research-backed interventions. CA system prompts should also avoid unnecessary repetition, wait for indications from the user to introduce solutions, and vary the structure and length of the messages to increase the similarity to human speech.

### 8.3 RQ3

This paper's last RQ aims to evaluate the developed experience: **How are CBT-based CA interventions informed by users' psychometrics perceived by lonely young adults?** The subquestions focus on different elements of the experience, including the effect and efficacy of the system prompts, the participants' opinions on the interaction, and the feasibility of using profiles to offer personalized CBT interventions. The data relevant to these questions was gathered during the user testing but is informed and complemented by the other two RQs. Again, no literature addresses this RQ precisely, partially due to its reliance on the previous research questions, but primarily because of the lack of research at this intersection.

**How effectively can CA system prompts apply the intended loneliness interventions?** In the evaluation of the commercial CA, Wya, Beatty et al.<sup>93</sup> found that the CBT methods attempted by the CA were effectively reproduced and were able to establish therapeutic bonds with the users. The interactions found in this research did apply different rates of conversational strategies depending on the system prompt. For example, Anne's interaction used more summarization than abstraction and more encouragement than any other interaction. When participants talked about the conversations they had with the two interactions, distinct components of the CBT interventions were described, but were not identified by

participants without context. After they were given the system prompts, participants recognized some of the strategies employed from their own experience with the interactions. Participants believed that any missing conversational strategies would have appeared with longer use sessions with the interactions.

**To what extent do different system prompts offer participants a different experience?** In the studies that evaluated personalization in the systematic review by Kocaballi et al.<sup>121</sup>, outcomes could not definitively be attributed to this component. To address this, they recommend structuring personalization around the theoretical framework developed by Fan and Poole<sup>130</sup> that explores what is personalized, to whom it is personalized, how explicit the personalization is, and the social goal of the personalization. According to this framework, the implementation in this research is an implicit, categorical, content-based, relational personalization, which is primarily connected to affect. This kind of personalization is subtle if it is not called out due to the lack of transparency (implicit), the non-individual grouping (categorical), and the emotional effect that may be hard to recognize. When considering this position in the framework, it is unsurprising that several participants were unable to notice that the interactions were distinct. When participants did notice, they tended to assume the differences were due to the details and character of the personas' stories.

In the evaluation, it was found that the participants who evaluated the Anne interaction were more critical than the participants who evaluated the Jessica interaction. However, the complicating factor when addressing this research question is participants only got to use each interaction with one situation. This means that they had little to truly compare because the differences in participants' opinions could be due to the persona's unique situation or participants' relation to the persona instead of the differences that came from the system prompt.

**Does the use of profiles offer adequate personalization to users?** Loneliness is rarely explored with psychometric profiles, and instead tends to be a factor traced within a network. The most similar research includes Käll's latent profile analysis, which classified people seeking help with loneliness into five profiles based on their psychometrics<sup>8</sup>. A tangential piece of literature is a study of the interactions between only emotional regulation and loneliness, which resulted in seven profiles; these tended to vary most by the frequency the individuals used maladaptive coping strategies<sup>131</sup>. No research outside of this thesis has been conducted that uses loneliness and other related psychometrics to inform mental health personalization, which makes Study 2's use of profiles relatively unique. However, this attempt was not met with great success; profiles were missing and the profile statements did not enable a less-invasive way of matching participants with profiles.

The primary way participants connected to profiles was through their ex-

ploration and adoption of personas. After prompted to find personas with the most similarities to them, some participants felt a strong connection to certain personas. However, these personas were rarely the most psychometrically similar to themselves. Instead, they tended to connect more with other characteristics of the personas, like gender, age, and geographic region, instead of the characteristics that related to the personas' experience with loneliness. This trend does align with new research about user modeling that indicates that the age and gender of users and the gender of the persona affect persona perceptions and the user-persona relation<sup>132</sup>. Further research has indicated that a persona's image, especially their smile, has a significant effect on perceived similarity and credibility<sup>7</sup>. The AI-generated images of neutral faces used for the profiles may have then been a detriment, and the small differences in smile could explain part of the difference in appeal between Anne (chosen four times) and Tanja (never chosen).

Even when participants did connect to psychometrically similar personas, only one participant chose the related interaction as their preferred interaction. It could be, as U1 stated, that the most-fitting interventions "*might be what I need to hear, but it's not what I wanted to hear.*" Based on the limited sample size and the number of profiles added during Study 2, it is safe to say that this research does not cover all the existing profiles within this population based on the psychometrics used. More research would need to be conducted to uncover profiles and allow for the design to fit all users.

To address this RQ subquestion, the profiles used in this research did not offer adequate personalization to users. However, it is also clear that profiles are not a monolith, and the participants within them differ in many elements of their lives, opinions, and desires. Additionally, while people may simply not like their prescribed intervention, the intervention fit may also be improperly applied. With increased population cover and verified connections between profiles and psychological interventions, though, profiles could be a valuable way to offer goal personalization reminiscent of mental health professionals.

**How do users perceive the CA and their interaction experiences?** The participants of Study 2 found the chat modality and CA's use of questions made the interaction easy. Participants thought the questions were relevant and promoted helpful reflection, and appreciated clarifying questions as they were seen as a way to improve the CA's accuracy. They felt that the interaction validated their interpretation of the situation, but that the CA sometimes focused on the wrong details. They felt the tone seemed professional, if maybe a bit formal, but some found the positive focus untrustworthy or fake. Users stated a preference for short, empathetic messages, which were perceived as more human. Some participants found certain interactions unnecessarily repetitive and believed that they would not be as effective as human conversations at restoring mood. In the quantitative evalu-



ation, participants found the interactions usable, dependable, worthwhile, novel, and easy to learn. However, the interaction had low engagement, poor stimulation, average aesthetics, high perceived effort, and did not teach the participants much. Although participants liked the interactions and found them valuable, many components could be improved to better meet user expectations and preferences.

**RQ3: How are CBT-based CA interventions informed by users' psychometrics perceived by lonely young adults?**

System prompts can provide users with similar experiences to psychological interventions. These interventions are unlikely to be recognized as pervasive strategies or interventions without being shared with the user. Instead, participants tended to attribute differences in interactions to the variation of content shared with them. The use of psychometric profiles as personalization has a strong potential as these psychometric characteristics are often foundational in the psychological understanding of loneliness and its interventions. However, success with this strategy depends on a full understanding of the psychometric landscape within the population and a confirmed connection between the profiles and the interventions. In the end, the participants tended to not match with or evaluate the intervention based on the most psychometrically similar profile. While this limits the ability to address this RQ, participants still found the interactions to be usable, validating, rewarding, and worthwhile.

## 8.4 Research Contribution

This research offers the research community and the mental health community several valuable contributions, primarily related to CA development and future research. Based on the qualitative research completed, there is now a better understanding of the lonely young adult population and their desires related to interactions with CAs. These studies offer insight into their preferred tones, conversation strategies, security standards, interaction modes, and amount of user control. Additionally, this research provides working CBT-based system prompts for LLaMa 3 8b, an open-source LLM model. This enables people to apply the benefits enjoyed by the participants easily. Finally, although it was not able to be adequately applied to the personalization of the CA interaction, this research's framework for profile development regarding loneliness and the identified profiles can be used as a foundation for other studies.

## 8.5 Limitations

The limitations of this research can be grouped into five topics related to the nature of human-computer interaction research and point to areas for future study.

**Technical issues, ephemerality, and interviewer influence:** The impact of the environment, improperly operating materials, and the reactions and presence of the researcher cannot be predicted or removed from the interpretation of the data. Two of the interviews were held in public cafes to accommodate participants who could not travel to Utrecht, which could have influenced the conversation topics and the participants' comfort levels. Additionally, out of the 14 participants, six encountered technical difficulties resulting in false starts, stilted conversations, or elongated testing sessions to manage system restarts. Further, our emotions and experiences change constantly, and having the same interview with a participant on another day could have resulted in drastically varied reactions and experiences. Finally, another interviewer would likely have brought up alternative topics in small talk, responded to participants' experiences uniquely, or had an identity to which participants related differently; these elements could have elicited further disclosure or prevented topics from coming up. Although this research includes all research protocols and the materials used, the recorded answers were uniquely tied to those specific moments and can not be reproduced.

**Lack of Deep Expertise:** This thesis was highly multidisciplinary, incorporating theoretical and applied psychology, psychometrics, qualitative human-computer interaction research, statistical methods, application development, and LLM customization and use. While adequately applied during this project's scope, these are all incredibly deep topic areas in which many academics and professionals spend their entire careers. With access to more expert feedback in any of these areas, this research may have incorporated different choices.

A strong psychological background and the use of more than one framework may have better informed the connections between the profiles and the assigned interventions. In the development of Study 2, many of the connections made between participants' opinions and experiences and the CBT strategies assigned to their profiles were based on nonprofessional interpretations and direct statements of preference, which may be lacking foundation or ignorant to some personal truths. Secondly, although an ethics-based goal for this research was to use a self-hosted LLM to best retain control over all participant data, the technical support required to do this was not readily available. The time constraints imposed on the research limited the scope of the technology and impacted the research design. Future research in this area would benefit from a stronger foundation of experience in these areas.

**Construct Validity:** The ability of a tool to accurately measure a concept is the construct validity, which can have a strong impact on research interpretation.

Although the psychometric tools and usability questionnaires tools have been verified in the cited research papers and work for large groups of people, there are individual influences that may have affected this research's results. Several participants asked about the meaning of some of the words or phrases due to lower English fluency levels. Others asked for boundaries and interpretations of many of the statements. U5 explained in their interview, "*difficulty in making eye contact can be very characteristic[ally] autistic, but does not have to mean that they are also socially anxious. [...] It can come from a different difficulty.*" The interpretations and explanations given were likely not identical to those which the original researchers would offer, which could have affected the tool's accuracy or precision. Specifically for UCLA-LS-9, different experiences of loneliness may not be accurately captured, and misinterpretations could have led to respondents' unnecessary exclusion from the research. Improved consistency in individual-level interpretations of the psychometrics used in this research could help make the use of these tools more consistent and reliable.

Separately from the psychometrics, the use of personas as a proxy for the profiles was a flawed strategy as users tended to select profiles that matched their gender identity, age, ethnicity, and other personal characteristics. Making these characteristics more ambiguous, or adjusting them based on the participant, may influence the personas they select and could result in choices that more closely resemble their psychometric profile. Additionally, the profile-based statements were more ambiguous than intended, and while the disconnect between the stated perspectives in the interviews and the ratings given in the survey could be due to a lack of self-awareness, it could have also been due to hyperbole or miscommunication on either side.

**Population and Sampling:** The population selected for this research was chosen thoughtfully, based primarily on a lack of existing research on the age range, factors about access, and an absence of mental health professionals involved, but an ideal population for this research would have had fewer exclusion criteria to offer a more comprehensive understanding of loneliness. Elements of the sampling may also limit how well findings can be extrapolated within this population as well. Although the population is already geographically defined, the recruiting methods likely narrowed the sampled population through the focus of fliers and recruiting materials on or in forums related to Utrecht's Universities. Additionally, only three of the 14 participants grew up in the Netherlands, which is a far lower rate of Dutch citizens than the sampled population; this disparity may have been due to the use of English.

**Participants and Testing:** Finally, elements of the research design limit the ability to extrapolate and color the interpretation of the results. This research only explored the experiences of 14 people, a small sample size that could not capture

the range and diversity of people in the target population. There must be absent or under-represented opinions and experiences, missing psychometric profiles, and more variation within profiles. Although the sample covers most of the age range, only three of the fourteen participants were above 30 years old; this younger skew likely had an impact on the results.

Related to the user testing design, the fact that one of the personas was never tested and two were only tested once offers little to no feedback on those personas and the connected interventions. Users may have enjoyed or benefited from those interactions but missed the opportunity because of a lack of connection with the persona, possibly for reasons unrelated to the profile (e.g., age, gender, country of origin). Several participants mentioned that using themselves as the 'main character' was more enjoyable and valuable than using the persona's situations. Further, participants only evaluated and explored one persona's interaction with personal experiences; the fact that this interaction was their preferred one means that the evaluations are likely more positive than more comprehensive evaluations of all the interactions. Finally, the participants were limited in how long they were able to use the interactions to keep the user testing sessions within a reasonable time frame. While this enabled testing and analysis in the constraints of this thesis, this research cannot extrapolate findings to apply to longer and repeated interactions with CAs in interventions about loneliness. Thus, allowing participants to use multiple interactions with their personal experiences for a less restricted time would have been more representative of their evaluations of the interactions.

## **8.6 Future Research**

The testing described in this paper was not an inappropriate way to gather qualitative information about the users' experiences. However, it may be inadequate for future research and testing with the goal of more vast extrapolation and more thorough evaluation.

Future research can utilize the structure of Study 1 and the interaction used in Study 2 on a larger scale. This research, with a less limited population regarding geographic location, mental health, and language, would enable the study of more diverse participants and provide a better picture of the help-seeking population. Exploring the experiences of a more diverse selection of lonely individuals, the perceptions and opinions that depend on certain user characteristics, and their needs and expectations related to a CA interaction would allow for better extrapolation. Another line of research could focus on a survey that gathers psychometric data along with experiences of loneliness and opinions about CA interaction qualities. An increased focus on quantitative data can better inform the creation of profiles and may allow for the validation of the profiles described in this research.

A valuable addition to this survey could be the identification of qualities that can discern the ideal CA personalization. A valuable addition to either of these studies would be the employment to experienced psychologists that could better pair the profiles or other psychological qualities with the best CBT strategy for that group.

Further removed studies could include an online version of the interaction and private, longer-term, repeated interactions with participants. This longitudinal study would offer a more realistic use context and enable the use of psychometrics as repeated outcome measures to infer efficacy. Additionally, as giving users a choice of system characteristics before using the interactions is a recommendation of this research, evaluating the choices that users make and how their decisions affect their experience is an important next step. More tangentially, the use of personas in this research led to the finding that people tend to connect more with personas that share geographic background, age, and gender with the participants. Exploring this dynamic and adapting personas to be closer to each participant's characteristics could lead to a better connection between persona and participant characteristics in the studied domain. Finally, changes to the interaction's design and features could result in new dynamics and areas to study. Voice interaction, as mentioned in interviews and suggested as a valuable addition, is a worthwhile area to consider. Existing research, including that of Kocielnik et al.<sup>133</sup>, asserts that voice interactions are better at prompting and supporting self-reflection. Future research can explore the differences in the voice interaction for lonely participants, especially those with more anxious symptoms who tended to be opposed to this mode. Another way the interaction could be updated is in the progressive adaptation of the system prompt based on data or characteristics gathered about the participant during the interaction. This task of eliciting relevant information, capturing the applicable components, transforming them into usable actions, and applying changes to the system prompt could allow for a highly personalized interaction.

## 9 Conclusion

Loneliness, while a pervasive and nearly universal experience, remains stigmatizing instead of unifying. Despite the high prevalence of loneliness among people under 40, there is a clear lack of research focusing on this demographic. The causes and relationships with loneliness of this age group tend to differ from both younger and older populations, instead more closely related to social comparisons and maladaptive cognitions. This disparity has limited the efficacy of the existing research and led to solutions that do not appeal to this demographic, support their needs, or address their concerns. More research aimed at supporting young adults who struggle with negative, destructive cognitions could lower

rates of mental distress and promote happier, healthier lives. The development of a scalable, adaptable, psychology-informed CA interaction for young adults to discuss their loneliness would be valuable to many people who would instead struggle alone.

This research demonstrates that members of the young adult population are interested in using a thoughtfully developed and effective CA interaction. Participants appreciated the CA's ability to reflect their thoughts, pose questions, and prompt self-reflection. The potential to encourage more people to explore their emotions related to loneliness is promising and could lead to improved mental health outcomes. As Fred Rogers of *Mister Rogers' Neighborhood* said, "*Anything that's human is mentionable, and anything that is mentionable can be more manageable. When we can talk about our feelings, they become less overwhelming, less upsetting, and less scary.*"<sup>134</sup> Fostering open, emotions-based conversations, especially about a stigmatized experience like loneliness, can help people in distress. LLMs can facilitate this process and the possibility is worth pursuing.

## 10 Acknowledgements

I would first like to thank the respondents to my screening survey and the participants who took part in the research. I could not have learned what I did without you offering your time and your mental energy, and I greatly appreciate the tiny insight I got into your lives and understanding of loneliness. I also could not have done this research without the guidance and encouragement of my supervisors, Hanna Hauptmann and Christof van Nimwegen; thank you for all you did. I also want to recognize some of my friends and peers who made the thesis process more possible and enjoyable. Thank you, Lukas, for your help with development and the support you gave in ideation. Thank you, Mark, for helping to pilot test my methods and prepare me for presentations. And thank you to Sarah, Roy, and the rest of the HCI Thesis Vibes group who were an excellent community of support, shared experiences, and communal stress. And a shout out to my friends who constantly heard me say I was busy with "thesis stuff"; thank you for the patience, feedback, and hype you offered. Finally, thank you Mom and Dad for giving me the opportunity to pursue this master's degree, for your belief in me every step of the way, and for raising me to become someone who trusts themselves enough to take on this monster of a project.

Finally, I'd like to thank the AI that helped me in this thesis. Here's to Word Online for doing some heavy lifting when it came to transcripts. To Bing AI for being my coding partner. To LLaMa 2 and LLaMa 3 who allowed me to vent under the guise of system prompt testing. And to Grammarly, for reminding me how to spell and telling me when I'm missing a preposition. Cheers!

## References

- [1] Roemer, A. & Medvedev, O. N. Positive and negative affect schedule (PANAS). In Medvedev, O. N., Krägeloh, C. U., Siegert, R. J. & Singh, N. N. (eds.) *Handbook of Assessment in Mindfulness Research*, 1–11 (Springer International Publishing). URL [https://doi.org/10.1007/978-3-030-77644-2\\_86-1](https://doi.org/10.1007/978-3-030-77644-2_86-1).
- [2] American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders* (American Psychiatric Association), fifth edition edn. URL <https://psychiatryonline.org/doi/book/10.1176/appi.books.9780890425596>.
- [3] AI@Meta. Llama 3 model card (2024). URL [https://github.com/meta-llama/llama3/blob/main/MODEL\\_CARD.md](https://github.com/meta-llama/llama3/blob/main/MODEL_CARD.md).
- [4] Cacioppo, J. T. & Hawkley, L. C. Perceived social isolation and cognition **13**, 447–454. URL [https://www.cell.com/trends/cognitive-sciences/abstract/S1364-6613\(09\)00147-8](https://www.cell.com/trends/cognitive-sciences/abstract/S1364-6613(09)00147-8). Publisher: Elsevier.
- [5] Käll, A. *et al.* A common elements approach to the development of a modular cognitive behavioral theory for chronic loneliness **88**, 269–282.
- [6] Eldridge, S. Percentile | definition, quartile, & facts. URL <https://www.britannica.com/topic/percentile>.
- [7] Salminen, J., Jung, S.-G., Santos, J. M. & Jansen, B. J. Does a smile matter if the person is not real?: The effect of a smile and stock photos on persona perceptions **36**, 568–590. URL <https://doi.org/10.1080/10447318.2019.1664068>. Publisher: Taylor & Francis eprint: <https://doi.org/10.1080/10447318.2019.1664068>.
- [8] Käll, A. Internet-based interventions for loneliness : Efficacy and latent psychopathological profiles of treatment seekers. ISBN: 9789179290573.
- [9] Veronese, N. *et al.* Interventions for reducing loneliness: An umbrella review of intervention studies **29**, e89–e96.
- [10] Society, P. What is psychometrics? URL <https://www.psychometricsociety.org/what-psychometrics>.
- [11] Le Blanc, A. L. *et al.* Evaluation of the psychometric properties of two short forms of the social interaction anxiety scale and the social phobia scale **21**,

- 312–323. URL <https://doi.org/10.1177/1073191114521279>. Publisher: SAGE Publications Inc.
- [12] Giray, L. Prompt engineering with ChatGPT: A guide for academic writers **51**, 2629–2633. URL <https://doi.org/10.1007/s10439-023-03272-4>.
- [13] Peplau, L. A. & Perlman, D. *Loneliness: A Sourcebook of Current Theory, Research and Therapy* (John Wiley & Sons, Inc.).
- [14] Barreto, M. *et al.* Exploring the nature and variation of the stigma associated with loneliness **39**, 2658. URL <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9386761/>. Publisher: SAGE Publications.
- [15] Murthy, V. Our epidemic of loneliness and isolation: The u.s. surgeon general’s advisory on the healing effects of social connection and community. URL <https://www.hhs.gov/sites/default/files/surgeon-general-social-connection-advisory.pdf>.
- [16] Frazer, L. & Andrew, S. Tackling loneliness annual report march 2023: the fourth year. URL <https://www.gov.uk/government/publications/loneliness-annual-report-the-fourth-year/tackling-loneliness-annual-report-march-2023-the-fourth-year>.
- [17] van Ooijen, M. Voortgangsrapportage één tegen eenzaamheid 2023. URL <https://open.overheid.nl/documenten/9f1b8d0f-9d70-4eb5-8062-33a470c75d97/file>.
- [18] Patel, V. *et al.* Reducing the treatment gap for mental disorders: a WPA survey **9**, 169–176. URL <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2953637/>.
- [19] Ngui, E. M., Khasakhala, L., Ndeti, D. & Roberts, L. W. Mental disorders, health inequalities and ethics: A global perspective **22**, 235–244. URL <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2935265/>.
- [20] Khawaja, Z. & Bélisle-Pipon, J.-C. Your robot therapist is not your therapist: understanding the role of AI-powered mental health chatbots **5**. URL <https://www.frontiersin.org/journals/digital-health/articles/10.3389/fdgth.2023.1278186>.
- [21] Sha, S. *et al.* Efficacy of relational agents for loneliness across age groups: a systematic review and meta-analysis **24**. URL <https://link.springer.com/article/10.1186/s12889-024-19153-x#article-info>.



- [22] Oh, Y. H., Chung, K. & Ju, D. Y. Differences in interactions with a conversational agent **17**, 3189. URL <https://www.mdpi.com/1660-4601/17/9/3189>. Number: 9 Publisher: Multidisciplinary Digital Publishing Institute.
- [23] Shah, H. A. & Househ, M. Understanding loneliness in younger people: Review of the opportunities and challenges for loneliness interventions **12**, e45197. URL <https://www.i-jmr.org/2023/1/e45197>. Company: Interactive Journal of Medical Research Distributor: Interactive Journal of Medical Research Institution: Interactive Journal of Medical Research Label: Interactive Journal of Medical Research Publisher: JMIR Publications Inc., Toronto, Canada.
- [24] Cacioppo, J. T., Hawkley, L. C. & Thisted, R. A. Perceived social isolation makes me sad: 5-year cross-lagged analyses of loneliness and depressive symptomatology in the Chicago Health, Aging, and Social Relations Study **25**, 453–463. Place: US Publisher: American Psychological Association.
- [25] Maes, M., Qualter, P., Lodder, G. M. A. & Mund, M. How (not) to measure loneliness: A review of the eight most commonly used scales **19**, 10816. URL <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9518421/>.
- [26] Weiss, R. S. *Loneliness: The Experience of Emotional and Social Isolation* (The Massachusetts Institute of Technology).
- [27] Ellard, K. K., Fairholme, C. P., Boisseau, C. L., Farchione, T. J. & Barlow, D. H. Unified protocol for the transdiagnostic treatment of emotional disorders: Protocol development and initial outcome data **17**, 88–101.
- [28] Heu, L. C. *et al.* Loneliness across cultures with different levels of social embeddedness: A qualitative study **28**, 379–405. URL <https://onlinelibrary.wiley.com/doi/abs/10.1111/pere.12367>. \_eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/pere.12367>.
- [29] Nicolaisen, M. & Thorsen, K. Who are lonely? Loneliness in different age groups (18–81 years old), using two measures of loneliness **78**, 229–257. URL <https://doi.org/10.2190/AG.78.3.b>. Publisher: SAGE Publications Inc.
- [30] Cacioppo, J. T., Hawkley, L. C., Norman, G. J. & Berntson, G. G. Social isolation **1231**, 17–22. URL <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1749-6632.2011.06028.x>. \_eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1749-6632.2011.06028.x>.

- [31] Luhmann, M. & Hawkley, L. C. Age differences in loneliness from late adolescence to oldest old age **52**, 943–959. URL <https://europepmc.org/articles/PMC8015413>.
- [32] O’Day, E. B. & Heimberg, R. G. Social media use, social anxiety, and loneliness: A systematic review **3**, 100070. URL <https://www.sciencedirect.com/science/article/pii/S245195882100018X>.
- [33] Ernst, M. *et al.* Loneliness before and during the COVID-19 pandemic: A systematic review with meta-analysis **77**, 660–677.
- [34] Lim, M. H., Eres, R. & Vasan, S. Understanding loneliness in the twenty-first century: an update on correlates, risk factors, and potential solutions **55**, 793–810. URL <https://doi.org/10.1007/s00127-020-01889-7>.
- [35] Surkalim, D. L. *et al.* The prevalence of loneliness across 113 countries: systematic review and meta-analysis **376**, e067068. URL <https://www.bmj.com/content/376/bmj-2021-067068>. Publisher: British Medical Journal Publishing Group Section: Research.
- [36] Statistiek, C. B. v. d. Eenzaamheid mensen van 15 jaar of ouder, 2019 t/m 2022. URL <https://www.cbs.nl/nl-nl/maatwerk/2023/38/eenzaamheid-mensen-van-15-jaar-of-ouder-2019-t-m-2022>. Last Modified: 2023-09-22T00:00:00+02:00.
- [37] Maes, M., Qualter, P., Vanhalst, J., Van den Noortgate, W. & Goossens, L. Gender differences in loneliness across the lifespan: A meta-analysis URL [https://journals-sagepub-com.proxy.library.uu.nl/doi/full/10.1002/per.2220?casa\\_token=DzTm-gSQvHsAAAAA%3AZHOPXInzkq6M70avueVi2Hpd0k\\_cJ1uREfc7wo8cf1c5XSL7djUU-Q0oaJ3WcCKQWpk4X1t15YkwQg](https://journals-sagepub-com.proxy.library.uu.nl/doi/full/10.1002/per.2220?casa_token=DzTm-gSQvHsAAAAA%3AZHOPXInzkq6M70avueVi2Hpd0k_cJ1uREfc7wo8cf1c5XSL7djUU-Q0oaJ3WcCKQWpk4X1t15YkwQg).
- [38] Mund, M. *et al.* Would the real loneliness please stand up? the validity of loneliness scores and the reliability of single-item scores **30**, 1226–1248. URL <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10149889/>.
- [39] Panayiotou, M., Badcock, J. C., Lim, M. H., Banissy, M. J. & Qualter, P. Measuring loneliness in different age groups: The measurement invariance of the UCLA loneliness scale **30**, 1688–1715. URL <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10248311/>.
- [40] House, J., Landis, K. & Umberson, D. Social relationships and health URL <https://www-science-org>.

proxy.library.uu.nl/doi/abs/10.1126/science.  
3399889?casa\_token=aH4calxR648AAAAA:U3CSByplqs2m\_  
9KfZzoRK8104R3THRZuB101TrFzXEKPXLlk50MinMm-W18gRZpby9FPVsFzio-jawU.

- [41] Hawkley, L. C. & Cacioppo, J. T. Aging and loneliness: Downhill quickly? URL <https://journals.sagepub.com/doi/full/10.1111/j.1467-8721.2007.00501.x>.
- [42] Hawkley, L. C. & Cacioppo, J. T. Loneliness matters: a theoretical and empirical review of consequences and mechanisms **40**, 218–227.
- [43] Wilson, R. S. *et al.* Loneliness and risk of alzheimer disease **64**, 234–240. URL <https://doi.org/10.1001/archpsyc.64.2.234>.
- [44] Cacioppo, J. T. *et al.* Lonely traits and concomitant physiological processes: the MacArthur social neuroscience studies **35**, 143–154. URL <https://www.sciencedirect.com/science/article/pii/S0167876099000495>.
- [45] Southward, M. W., Terrill, D. R. & Sauer-Zavala, S. The effects of the unified protocol and unified protocol skills on loneliness in the COVID-19 pandemic **39**, 913–921.
- [46] Teneva, N. & Lemay, E. P. Projecting loneliness into the past and future: implications for self-esteem and affect **44**, 772–784. URL <https://doi.org/10.1007/s11031-020-09842-6>.
- [47] Gow, A. J., Pattie, A., Whiteman, M. C., Whalley, L. J. & Deary, I. J. Social support and successful aging **28**, 103–115. URL <https://econtent-hogrefe-com.proxy.library.uu.nl/doi/abs/10.1027/1614-0001.28.3.103>. Publisher: Hogrefe Publishing.
- [48] Van Beljouw, I. M., Verhaak, P. F., Cuijpers, P., Van Marwijk, H. W. & Penninx, B. W. The course of untreated anxiety and depression, and determinants of poor one-year outcome: a one-year cohort study **10**, 86. URL <https://bmcp psychiatry.biomedcentral.com/articles/10.1186/1471-244X-10-86>.
- [49] McHugh Power, J. *et al.* Depressive symptoms predict increased social and emotional loneliness in older adults **24**, 110–118. URL <https://doi.org/10.1080/13607863.2018.1517728>. Publisher: Routledge eprint: <https://doi.org/10.1080/13607863.2018.1517728>.

- [50] Lim, M. H., Rodebaugh, T. L., Zyphur, M. J. & Gleeson, J. F. M. Loneliness over time: The crucial role of social anxiety **125**, 620–630. URL <https://oce-ovid-com.proxy.library.uu.nl/article/00004468-201607000-00002/HTML>.
- [51] Hawkley, L. C., Thisted, R. A. & Cacioppo, J. T. Loneliness predicts reduced physical activity: Cross-sectional & longitudinal analyses **28**, 354–363. Place: US Publisher: American Psychological Association.
- [52] Baumeister, R. F., DeWall, C. N., Ciarocco, N. J. & Twenge, J. M. Social exclusion impairs self-regulation. **88**, 589–604. URL <http://doi.apa.org/getdoi.cfm?doi=10.1037/0022-3514.88.4.589>.
- [53] Penedo, F. J. & Dahn, J. R. Exercise and well-being: a review of mental and physical health benefits associated with physical activity **18**, 189–193.
- [54] Dahlberg, L. & Krug, E. The social-ecological model: A framework for prevention. URL <https://www.cdc.gov/violenceprevention/about/social-ecologicalmodel.html>.
- [55] Käll, A. Request to learn about internet-based interventions for loneliness.
- [56] VanderWeele, T. J., Hawkley, L. C., Thisted, R. A. & Cacioppo, J. T. A marginal structural model analysis for loneliness: Implications for intervention trials and clinical practice **79**, 225–235. Place: US Publisher: American Psychological Association.
- [57] McWhirter, B. T. Loneliness: A review of current literature, with implications for counseling and research **68**, 417. URL <https://search.ebscohost.com/login.aspx?direct=true&db=buh&AN=5343737&site=ehost-live>. Publisher: Wiley-Blackwell.
- [58] Masi, C. M., Chen, H.-Y., Hawkley, L. C. & Cacioppo, J. T. A meta-analysis of interventions to reduce loneliness **15**, 219–266. URL [https://journals-sagepub-com.proxy.library.uu.nl/doi/abs/10.1177/1088868310377394?casa\\_token=dEegGfQAsz4AAAAA:BaHXuyWyMP13zF1mbZP0o730UpTGkSSS2XoVZhZD8ZerrCsEG1P1Mxtn3iIuLFA-HdXMijQGtSOQ](https://journals-sagepub-com.proxy.library.uu.nl/doi/abs/10.1177/1088868310377394?casa_token=dEegGfQAsz4AAAAA:BaHXuyWyMP13zF1mbZP0o730UpTGkSSS2XoVZhZD8ZerrCsEG1P1Mxtn3iIuLFA-HdXMijQGtSOQ).
- [59] Erhardt, D., Bunyi, J., Dodge-Rice, Z., Neary, M. & Schueller, S. M. Digitized thought records: a practitioner-focused review of cognitive restructuring apps **15**, e39. URL <https://www.cambridge.org/core/journals/the-cognitive-behaviour-therapist/article/digitized-thought-records-a-practitionerfocused-review-of-cognitive-restru>

7D79B49EEF560F78E1534F5C6DA264CD. Publisher: Cambridge University Press.

- [60] Frank, J. D. & Frank, J. B. *Persuasion and healing: A comparative study of psychotherapy, 3rd ed.* Persuasion and healing: A comparative study of psychotherapy, 3rd ed (Johns Hopkins University Press). Pages: xviii, 343.
- [61] Brown, J. E. H. & Halpern, J. AI chatbots cannot replace human interactions in the pursuit of more inclusive mental healthcare **1**, 100017. URL <https://www.sciencedirect.com/science/article/pii/S2666560321000177>.
- [62] Bordin, E. S. The generalizability of the psychoanalytic concept of the working alliance **16**, 252–260. Place: US Publisher: Division of Psychotherapy (29), American Psychological Association.
- [63] Pennebaker, J., Zech, E. & Rimé, B. Disclosing and sharing emotion: Psychological, social, and health consequences 517–539.
- [64] Weiss, R. D., Jaffee, W. B., de Menil, V. P. & Cogley, C. B. Group therapy for substance use disorders: What do we know? **12**, 339. URL [https://journals.lww.com/hrpjournal/fulltext/2004/11000/group\\_therapy\\_for\\_substance\\_use\\_disorders\\_\\_what\\_do.4.aspx?casa\\_token=9dcfxgHDqCUAAAAA:htRnLdnF3nSSU6WH5wcLEmJCfBBHdR2Pi6e6WhU6PHRXq3coMOfwx0hSbGEEhHFSxcYiMjHfCEd](https://journals.lww.com/hrpjournal/fulltext/2004/11000/group_therapy_for_substance_use_disorders__what_do.4.aspx?casa_token=9dcfxgHDqCUAAAAA:htRnLdnF3nSSU6WH5wcLEmJCfBBHdR2Pi6e6WhU6PHRXq3coMOfwx0hSbGEEhHFSxcYiMjHfCEd)
- [65] Smyth, J. M. Written emotional expression: effect sizes, outcome types, and moderating variables **66**, 174–184.
- [66] Wenzel, A. Basic strategies of cognitive behavioral therapy **40**, 597–609. URL [https://www.psych.theclinics.com/article/S0193-953X\(17\)30067-9/fulltext](https://www.psych.theclinics.com/article/S0193-953X(17)30067-9/fulltext). Publisher: Elsevier.
- [67] Allen, L. B., McHugh, R. K. & Barlow, D. H. Emotional disorders: A unified protocol. In *Clinical Handbook of Psychological Disorders*, 216–250 (The Guilford Press), 4 edn.
- [68] Oti, O. & Pitt, I. Online mental health interventions designed for students in higher education: A user-centered perspective **26**, 100468. URL <https://www-sciencedirect-com.proxy.library.uu.nl/science/article/pii/S2214782921001081>. Publisher: Elsevier.

- [69] Muroff, J. & Robinson, W. Tools of engagement: Practical considerations for utilizing technology-based tools in CBT practice **29**, 81–96. URL <https://www.sciencedirect.com/science/article/pii/S1077722920300249>.
- [70] Chowdhary, K. R. Natural language processing. In Chowdhary, K. (ed.) *Fundamentals of Artificial Intelligence*, 603–649 (Springer India). URL [https://doi.org/10.1007/978-81-322-3972-7\\_19](https://doi.org/10.1007/978-81-322-3972-7_19).
- [71] Chang, Y. *et al.* A survey on evaluation of large language models URL <https://dl.acm.org/doi/10.1145/3641289>. Just Accepted.
- [72] Wu, T., Terry, M. & Cai, C. J. AI chains: Transparent and controllable human-AI interaction by chaining large language model prompts. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, CHI '22, 1–22 (Association for Computing Machinery). URL <https://dl.acm.org/doi/10.1145/3491102.3517582>.
- [73] Liu, Y. *et al.* Summary of ChatGPT-related research and perspective towards the future of large language models **1**, 100017. URL <https://www.sciencedirect.com/science/article/pii/S2950162823000176>.
- [74] Adamopoulou, E. & Moussiades, L. An overview of chatbot technology **584**, 373–383. URL <https://europepmc.org/articles/PMC7256567>.
- [75] Mott, B., Lester, J. & Branting, K. Conversational agents **20042960**. URL <http://www.crcnetbase.com/doi/abs/10.1201/9780203507223.ch10>. Book Title: *The Practical Handbook of Internet Computing* ISBN: 9781584883814 9781466526907 Publisher: Chapman and Hall/CRC.
- [76] Deng, Y., Lei, W., Huang, M. & Chua, T.-S. Rethinking conversational agents in the era of LLMs: Proactivity, non-collaborativity, and beyond. In *Proceedings of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval in the Asia Pacific Region*, SIGIR-AP '23, 298–301 (Association for Computing Machinery). URL <https://dl.acm.org/doi/10.1145/3624918.3629548>.
- [77] Friedman, L. *et al.* Leveraging large language models in conversational recommender systems. URL <http://arxiv.org/abs/2305.07961>. 2305.07961[cs].
- [78] Grodniewicz, J. P. & Hohol, M. Waiting for a digital therapist: three challenges on the path to psychotherapy delivered by artificial intelligence

14. URL <https://www.frontiersin.org/journals/psychiatry/articles/10.3389/fpsy.2023.1190084>.
- [79] Montagna, S., Ferretti, S., Klopfenstein, L. C., Florio, A. & Pengo, M. F. Data decentralisation of LLM-based chatbot systems in chronic disease self-management. In *Proceedings of the 2023 ACM Conference on Information Technology for Social Good, GoodIT '23*, 205–212 (Association for Computing Machinery). URL <https://dl.acm.org/doi/10.1145/3582515.3609536>.
- [80] Abd-alrazaq, A. A. *et al.* An overview of the features of chatbots in mental health: A scoping review **132**, 103978. URL <https://www.sciencedirect.com/science/article/pii/S1386505619307166>.
- [81] Zemčík, T. A brief history of chatbots .
- [82] Song, I., Pendse, S. R., Kumar, N. & De Choudhury, M. The typing cure: Experiences with large language model chatbots for mental health support. URL <http://arxiv.org/abs/2401.14362>. 2401.14362[cs].
- [83] Vaidyam, A. N., Wisniewski, H., Halamka, J. D., Kashavan, M. S. & Torous, J. B. Chatbots and conversational agents in mental health: A review of the psychiatric landscape **64**, 456–464. URL <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6610568/>.
- [84] Ferrera, A. Empowering emotional well-being through a LLM-based chatbot : a comparative study with the standard journaling technique. URL <https://www.politesi.polimi.it/handle/10589/209293>.
- [85] Izumi, K. *et al.* Response generation for cognitive behavioral therapy with large language models: Comparative study with socratic questioning URL <http://arxiv.org/abs/2401.15966>. 2401.15966[cs].
- [86] Glenn, T. & Monteith, S. Privacy in the digital world: Medical and health data outside of HIPAA protections **16**, 494. URL <https://doi.org/10.1007/s11920-014-0494-4>.
- [87] Van Drongelen, A., De Bruijn, A., Roszek, B. & Vonk, R. Apps under the medical devices legislation URL <https://rivm.openrepository.com/handle/10029/622397>. Publisher: RIVM.
- [88] Health. URL [https://www.edps.europa.eu/data-protection/our-work/subjects/health\\_en](https://www.edps.europa.eu/data-protection/our-work/subjects/health_en).

- [89] Hacker, P. A legal framework for AI training data—from first principles to the artificial intelligence act **13**, 257–301. URL <https://doi.org/10.1080/17579961.2021.1977219>. Publisher: Routledge eprint: <https://doi.org/10.1080/17579961.2021.1977219>.
- [90] Zhang, Z. *et al.* “it’s a fair game”, or is it? examining how users navigate disclosure risks and benefits when using LLM-based conversational agents. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*, CHI ’24, 1–26 (Association for Computing Machinery). URL <https://doi.org/10.1145/3613904.3642385>.
- [91] Denecke, K., Abd-Alrazaq, A. & Househ, M. Artificial intelligence for chatbots in mental health: Opportunities and challenges. In Househ, M., Borycki, E. & Kushniruk, A. (eds.) *Multiple Perspectives on Artificial Intelligence in Healthcare: Opportunities and Challenges*, Lecture Notes in Bioengineering, 115–128 (Springer International Publishing). URL [https://doi.org/10.1007/978-3-030-67303-1\\_10](https://doi.org/10.1007/978-3-030-67303-1_10).
- [92] Wysa - everyday mental health. URL <https://www.wysa.com/>.
- [93] Beatty, C., Malik, T., Meheli, S. & Sinha, C. Evaluating the therapeutic alliance with a free-text CBT conversational agent (wysa): A mixed-methods study **4**. URL <https://www.frontiersin.org/journals/digital-health/articles/10.3389/fdgth.2022.847991>.
- [94] Durden, E. *et al.* Changes in stress, burnout, and resilience associated with an 8-week intervention with relational agent “woeBot” **33**, 100637. URL <https://www.sciencedirect.com/science/article/pii/S2214782923000374>.
- [95] Jo, E., Epstein, D. A., Jung, H. & Kim, Y.-H. Understanding the benefits and challenges of deploying conversational AI leveraging large language models for public health intervention. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, CHI ’23, 1–16 (Association for Computing Machinery). URL <https://dl.acm.org/doi/10.1145/3544548.3581503>.
- [96] Dosovitsky, G. & Bunge, E. L. Bonding with bot: User feedback on a chatbot for social isolation **3**. URL <https://www.frontiersin.org/journals/digital-health/articles/10.3389/fdgth.2021.735053>.



- [97] Lyon, A. R. & Koerner, K. User-centered design for psychosocial intervention development and implementation **23**, 180–200. Place: United Kingdom Publisher: Wiley-Blackwell Publishing Ltd.
- [98] Dimock, M. Defining generations: Where millennials end and generation z begins. URL <https://www.pewresearch.org/short-reads/2019/01/17/where-millennials-end-and-generation-z-begins/>.
- [99] Wittenborg, J. The netherlands is top english-speaking country. URL <https://www.wittenborg.eu/netherlands-top-english-speaking-country.htm>.
- [100] Castaneda, A. E., Tuulio-Henriksson, A., Marttunen, M., Suvisaari, J. & Lönnqvist, J. A review on cognitive impairments in depressive and anxiety disorders with a focus on young adults **106**, 1–27. URL <https://www.sciencedirect.com/science/article/pii/S0165032707002297>.
- [101] Borghouts, J. *et al.* Barriers to and facilitators of user engagement with digital mental health interventions: Systematic review **23**, e24387. URL <https://www.jmir.org/2021/3/e24387>.
- [102] Pyle, E. & Evans, D. Loneliness - what characteristics and circumstances are associated with feeling lonely? 1–18.
- [103] Kroenke, K., Spitzer, R. L. & Williams, J. B. The PHQ-9: validity of a brief depression severity measure **16**, 606–613.
- [104] Spitzer, R. L., Kroenke, K., Williams, J. B. W. & Löwe, B. A brief measure for assessing generalized anxiety disorder: The GAD-7 **166**, 1092–1097. URL <https://doi.org/10.1001/archinte.166.10.1092>.
- [105] Peters, L., Sunderland, M., Andrews, G., Rapee, R. M. & Mattick, R. P. Development of a short form social interaction anxiety (SIAS) and social phobia scale (SPS) using nonparametric item response theory: The SIAS-6 and the SPS-6. **24**, 66–76. Place: US Publisher: American Psychological Association.
- [106] Crawford, J. R. & Henry, J. D. The positive and negative affect schedule (PANAS): Construct validity, measurement properties and normative data in a large non-clinical sample **43**, 245–265. URL <https://onlinelibrary.wiley.com/doi/abs/10.1348/0144665031752934>. eprint: <https://bpspsychub.onlinelibrary.wiley.com/doi/pdf/10.1348/0144665031752934>.

- [107] Jensen, M. B., Elverum, C. W. & Steinert, M. Eliciting unknown unknowns with prototypes: Introducing prototrials and prototrial-driven cultures **49**, 1–31. URL <https://www.sciencedirect.com/science/article/pii/S0142694X16300849>.
- [108] Sedgwick, P. The normal distribution **345**, e6533. URL <https://www.bmj.com/content/345/bmj.e6533>. Publisher: British Medical Journal Publishing Group Section: Endgames.
- [109] Byrd-Bredbenner, C., Eck, K. & Quick, V. GAD-7, GAD-2, and GAD-mini: Psychometric properties and norms of university students in the united states **69**, 61–66. URL <https://www.sciencedirect.com/science/article/pii/S0163834321000025>.
- [110] Bond, F. W. & Dryden, W. *Handbook of Brief Cognitive Behaviour Therapy* (John Wiley & Sons, Incorporated). URL <http://ebookcentral.proquest.com/lib/uunl/detail.action?docID=220527>.
- [111] replicate/llama-chat. URL <https://github.com/replicate/llama-chat>. Original-date: 2023-07-17T19:00:13Z.
- [112] Chat with meta llama 3 on replicate. URL <https://www.llama2.ai/>.
- [113] Microsoft. Bing copilot: An ai assistant (2024). Unpublished AI model by Microsoft.
- [114] Image creator from microsoft designer - bing. URL <https://www.bing.com/images/create>.
- [115] O'Brien, H. L., Cairns, P. & Hall, M. A practical approach to measuring user engagement with the refined user engagement scale (UES) and new UES short form **112**, 28–39. URL <https://www.sciencedirect.com/science/article/pii/S1071581918300041>.
- [116] Di Malta, G. *et al.* The patient-perceived helpfulness of measures scale: Development and validation of a scale to assess the helpfulness of using measures in psychological treatment 10731911231195837. URL <https://doi.org/10.1177/10731911231195837>. Publisher: SAGE Publications Inc.
- [117] Schrepp, M., Hinderks, A. & Thomaschewski, J. Design and evaluation of a short version of the user experience questionnaire (UEQ-s) **4**, 103–108. URL <https://idus.us.es/handle/11441/107084>. Accepted: 2021-04-14T11:12:39Z Publisher: UNIR.

- [118] Schrepp, M. An analysis of standard deviations for UEQ scales. URL <https://rgdoi.net/10.13140/RG.2.2.34969.08809>.
- [119] Wester, J., de Jong, S., Pohl, H. & van Berkel, N. Exploring people's perceptions of LLM-generated advice **2**, 100072. URL <https://www.sciencedirect.com/science/article/pii/S294988212400032X>.
- [120] Swift, J. K., Callahan, J. L., Cooper, M. & Parkin, S. R. The impact of accommodating client preference in psychotherapy: A meta-analysis **74**, 1924–1937. URL <https://onlinelibrary-wiley-com.proxy.library.uu.nl/doi/abs/10.1002/jclp.22680>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/jclp.22680>.
- [121] Kocaballi, A. B. *et al.* The personalization of conversational agents in health care: Systematic review **21**, e15360. URL <https://www.jmir.org/2019/11/e15360>. Company: Journal of Medical Internet Research Distributor: Journal of Medical Internet Research Institution: Journal of Medical Internet Research Label: Journal of Medical Internet Research Publisher: JMIR Publications Inc., Toronto, Canada.
- [122] Peräkylä, A. Conversation analysis and psychotherapy: Identifying transformative sequences **52**, 257–280. URL <https://doi.org/10.1080/08351813.2019.1631044>. Publisher: Routledge eprint: <https://doi.org/10.1080/08351813.2019.1631044>.
- [123] Ta, V. *et al.* User experiences of social support from companion chatbots in everyday contexts: Thematic analysis **22**, e16235. URL <https://www.jmir.org/2020/3/e16235>. Company: Journal of Medical Internet Research Distributor: Journal of Medical Internet Research Institution: Journal of Medical Internet Research Label: Journal of Medical Internet Research Publisher: JMIR Publications Inc., Toronto, Canada.
- [124] Ma, Z., Mei, Y. & Su, Z. Understanding the benefits and challenges of using large language model-based conversational agents for mental well-being support **2023**, 1105–1114. URL <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10785945/>.
- [125] Morris, R. R., Kouddous, K., Kshirsagar, R. & Schueller, S. M. Towards an artificially empathic conversational agent for mental health applications: System design and user perceptions **20**, e10148. URL <https://www.jmir.org/2018/6/e10148>. Company: Journal of Medical Internet Research Distributor: Journal of Medical Internet Research Institution:

Journal of Medical Internet Research Label: Journal of Medical Internet Research  
Publisher: JMIR Publications Inc., Toronto, Canada.

- [126] Card, K. G. & Skakoon-Sparling, S. Are social support, loneliness, and social connection differentially associated with happiness across levels of introversion-extraversion? **10**, 20551029231184034. URL <https://doi.org/10.1177/20551029231184034>. Publisher: SAGE Publications Ltd.
- [127] Lee, C.-Y. S. & Goldstein, S. E. Loneliness, stress, and social support in young adulthood: Does the source of support matter? **45**, 568–580. URL <http://link.springer.com/10.1007/s10964-015-0395-9>.
- [128] Yu, H. & McGuinness, S. An experimental study of integrating fine-tuned LLMs and prompts for enhancing mental health support chatbot system URL <https://repository.derby.ac.uk/item/q5140/an-experimental-study-of-integrating-fine-tuned-llms-and-prompts-for-enhancing-mental-health-support-chatbot-system>  
Accepted: 2024-03-07 Company: AME Publishing Company Distributor: AME Publishing Company Institution: AME Publishing Company Label: AME Publishing Company Publisher: 929v9.
- [129] Xu, X. *et al.* Mental-LLM: Leveraging large language models for mental health prediction via online text data **8**, 1–32. URL <http://arxiv.org/abs/2307.14385>. 2307.14385[cs].
- [130] Fan, H. & Poole, M. S. What is personalization? perspectives on the design and implementation of personalization in information systems **16**, 179–202. URL <https://doi.org/10.1080/10919392.2006.9681199>. Publisher: Taylor & Francis eprint: <https://doi.org/10.1080/10919392.2006.9681199>.
- [131] Tan, A. J. *et al.* Loneliness versus distress: A comparison of emotion regulation profiles **39**, 180–190. URL <https://www-cambridge-org.utrechtuniversity.idm.oclc.org/core/journals/behaviour-change/article/loneliness-versus-distress-a-comparison-of-emotion-regulation-profiles/E182E471E0FAAA7C41DDD70F0E3C5CC3>.
- [132] Kaate, I. *et al.* Modeling the new modalities of personas: How do users' attributes influence their perceptions and use of interactive personas? In *Adjunct Proceedings of the 32nd ACM Conference on User Modeling, Adaptation and Personalization*, UMAP Adjunct '24, 164–169 (Association for Computing Machinery). URL <https://doi.org/10.1145/3631700.3664882>.

- [133] Kocielnik, R., Xiao, L., Avrahami, D. & Hsieh, G. Reflection companion: A conversational system for engaging users in reflection on physical activity **2**, 70:1–70:26. URL <https://doi.org/10.1145/3214273>.
- [134] Harrington, N. G. & Record, R. A. *Health Communication: Research and Practice for a Diverse and Changing World* (Taylor & Francis). Google-Books-ID: JVHGEAAAQBAJ.

## A Recruitment Poster

**ARE YOU LONELY?**  
Consider participating in this research project

The goal is to develop a digital, conversational interaction to help lonely people based on their needs, context, and individual experience. No technical knowledge needed.

Interested? Complete the survey linked to this QR code or at this link:  
[http://bit.ly/loneliness\\_research](http://bit.ly/loneliness_research)

Currently recruiting for:  
**Interviews**

1.5 to 2 hours  
Snacks provided  
Aged 18 - 40  
In & near Utrecht  
Comfortable with English

Utrecht University  
Sharing science, shaping tomorrow

Research is conducted for the Human Computer Interaction programme of the Information and Computing Sciences department. Questions? For more info, email: [L.a.beehler@students.uu.nl](mailto:L.a.beehler@students.uu.nl)

*Figure 15. A landscape-oriented version of the recruitment poster for interviews.*

**ARE YOU LONELY?**

Participate in this research project

I've designed an AI chat interaction to support people dealing with loneliness. I aim to adapt this interactions based on user's needs, context, and other experiences. I'll explore your expectations, opinions, and suggestions with surveys, prototypes, and open interview questions. No technical knowledge needed.

Aged 18 - 40    In & near Utrecht    Comfortable with English    Approx. 1 hour    Snacks provided

Currently recruiting for:  
**Prototype Testing**

Interested? Learn more and complete the screening survey at this QR code or at this link:  
[http://bit.ly/loneliness\\_research](http://bit.ly/loneliness_research)

Research is conducted for the Human Computer Interaction programme of Utrecht University's Information and Computing Sciences department. Questions? For more information, email: [la.bechler@students.uu.nl](mailto:la.bechler@students.uu.nl)

Figure 16. A portrait-oriented version of the recruitment poster used for user testing.

## B Screening Survey Questions

This survey has been developed for the Human Computer Interaction master's thesis of Lauren Beehler. This thesis is focused on situational and chronic loneliness in people between the ages of 18 and 40. The aim is to develop an interaction to help people who are lonelier than average manage, accept, or decrease their loneliness experiences. This survey will discern if your characteristics fit within the target population of this thesis.

Please read the statements below and tick the final box to confirm you have read and understood the statements and upon doing so agree to participate in the project.

- I confirm that the research project has been explained to me. I have had the opportunity to ask questions about the project and have had these answered satisfactorily. I had enough time to consider whether to participate.
- I consent to the material I contribute being used to generate insights for the research project.
- I understand that personal data will be collected from me and that this information will be held confidentially so that only Lauren Beehler will have access to this data and can trace the information back to me personally. The information will be held in a password-protected drive for up to six months after which period it will be fully anonymised. In accordance with the General Data Protection Regulation (GDPR) I can have access to my information and can request my data to be deleted at any time during this period.
- I understand that my participation in this research is voluntary and that I may withdraw from the study at any time without providing a reason, and that if I withdraw any personal data already collected from me will be erased.
- I consent to allow the fully anonymized data to be used in future publications and other scholarly means of disseminating the findings from the research project.
- I understand that the data acquired will be securely stored by researchers, but that appropriately anonymized data may in the future be made available to others for research purposes.
- I understand that I can request any personal data collected from me to be deleted.



1. I confirm that I have read and understood the above statements, and agree to participate in the study.
  - (a) I confirm that I have read and understood the above statements, and agree to participate in the study.
  - (b) I do not agree to participate in this study.
2. Please enter the year you were born: \_\_\_\_
3. Please share your gender identity:
  - (a) Man
  - (b) Woman
  - (c) Non-binary
  - (d) Prefer not to say
4. What is your English fluency level?
  - (a) A1/A2 (Pre-Intermediate)
  - (b) B1 (Intermediate)
  - (c) B2 (Upper Intermediate)
  - (d) C1 (Advanced)
  - (e) C2 (Proficient)
  - (f) Native language
  - (g) I don't know
5. How often do you feel lonely?
  - (a) Often
  - (b) Sometimes
  - (c) Occasionally
  - (d) Hardly ever
  - (e) Never
6. Have you ever been diagnosed by a mental health professional with a psychiatric disorder or difference? (E.g., depression, general anxiety disorder, social anxiety disorder, PTSD, panic disorder, agoraphobia, eating disorder, alcohol use disorder)
  - (a) No

- (b) No, but I believe I could be diagnosed with: ----
  - (c) Yes, I have been diagnosed with: ----
7. Are you currently in psychiatric treatment?
- (a) No
  - (b) Yes, for the diagnosis mentioned above
  - (c) Yes, but for another concern
8. Have you started, stopped, or changed the dosage of a prescribed psychiatric medication in the last three months?
- (a) Yes
  - (b) No
9. Have you experienced any of these in the last month: (Select all that apply)
- (a) Suicidal thoughts or plans
  - (b) Serious thoughts of hurting others
  - (c) Substance abuse (use of alcohol or a drug in amounts or by methods that are harmful to you or others)
  - (d) None of the above
10. **[UCLA-LS-9]** How often do you feel that you lack companionship?
- (a) Never
  - (b) Rarely
  - (c) Sometimes
  - (d) Often
11. **[UCLA-LS-9]** How often do you feel that you are 'in tune' with the people around you?
- (a) Never
  - (b) Rarely
  - (c) Sometimes
  - (d) Often
12. **[UCLA-LS-9]** How often do you feel that there are people you can talk to?

- (a) Never
- (b) Rarely
- (c) Sometimes
- (d) Often

13. [UCLA-LS-9] How often do you feel left out?

- (a) Never
- (b) Rarely
- (c) Sometimes
- (d) Often

14. [UCLA-LS-9] How often do you feel that there are people you can turn to?

- (a) Never
- (b) Rarely
- (c) Sometimes
- (d) Often

15. [UCLA-LS-9] How often do you feel that you have a lot in common with the people around you?

- (a) Never
- (b) Rarely
- (c) Sometimes
- (d) Often

16. [UCLA-LS-9] How often do you feel close to people?

- (a) Never
- (b) Rarely
- (c) Sometimes
- (d) Often

17. [UCLA-LS-9] How often do you feel isolated from others?

- (a) Never
- (b) Rarely

- (c) Sometimes
- (d) Often

18. [UCLA-LS-9] How often do you feel part of a group of friends?

- (a) Never
- (b) Rarely
- (c) Sometimes
- (d) Often

19. Thank you for filling out the questions! You may qualify for this study; please leave your email here: -----

20. Please let me know if you have anything you want to share. This can include scheduling constraints, questions or clarifications about what was asked or the answers given, etc.: -----

## C Rejection & Mental Health Resources Email

Hello [Name],

My name is Lauren, and I want to thank you for recently filling in my survey about loneliness. Unfortunately, you don't appear to fit the population I'm researching. I know that an email like this can be disappointing, but I genuinely appreciate your interest.

Since my research is about loneliness, I've learned a lot about the concept beyond my personal experiences. I initially didn't realize how universal the experience of loneliness is; loneliness affects people across different backgrounds, ages, and life circumstances. I've learned that loneliness is the perception that you don't have your desired relationships and social connections and that loneliness can be a natural response to life events, transitions, or changes like moving to a new city, experiencing a loss, and graduating. I've learned that loneliness can be connected to depression, general anxiety, social anxiety, and several other mental health concerns.

However, I already knew that loneliness can hurt, and I knew that it takes courage to recognize and talk about feelings of loneliness. It's essential to recognize that loneliness is not a sign of weakness or failure. There is often a stigma associated with loneliness, which can prevent people from seeking support or talking with others about their experiences. You are not alone in this, and many others share similar emotions; that's the entire point of my research.

I encourage you to consider reaching out to someone—a friend, family member, or a mental health professional. If you want some additional support, here is a list of resources that you can turn to:

- Alles Oké? (Everything Ok?) – made for people between 18 and 25, you can chat or call at 0800-0450 between 14:00 and 22:00 in Dutch or English. The free and anonymous calls to their trained volunteers can touch on any topic.
- De Luisterlijn – a primarily Dutch-language program that allows people to anonymously call, text, or email at any time of day. Call 088-0767-000 to talk to a trained volunteer. The people of the Listening Line only give advice if requested, and mainly act as someone to talk to. Want praten lucht op.
- MIND Korrelatie – offers anonymous, professional, psychological, and psychosocial support to anyone who asks for it. This can be done by phone (call 0900-1450 or WhatsApp 06-13-86-38-03), via email (vragen@wijzijnmind.nl) or online. They have specific resources for those with depression symptoms and addiction.

- Caring Universities – if you're a student, Caring Universities offers free online services to improve student's mental well-being. They offer 9 free online programs and a survey that will help you decide which might be the most effective based on your current mood.
- UU Student Psychologist – if you're a student of Utrecht University, consider making an appointment with a student psychologist; they can help with short-term study problems and/or mild personal problems. The intake appointment costs €25 but further appointments are free of charge. If that doesn't sound like a good fit, you can also talk to a member of A Bright Conversation, a diverse group of facilitators interested in chatting with you about challenging life situations.
- Don't have a GP? Here's a website to find options near you: [ZorgkaartNederland](http://ZorgkaartNederland.nl). For students of Utrecht, here are some more resources for mental and physical health: [Studentenzorgwijzer.nl](http://Studentenzorgwijzer.nl).
- Find a Helpline – lists helplines in the Netherlands and other countries. Beyond those I already listed, it suggests the Center for Sexual Violence, Victim Support NL (offers emotional and practical support after intense events), Veilig Thuis (tip and support line for domestic violence), [99gram.nl](http://99gram.nl) (help for eating disorders), De Oudertelefoon (parenting support), and Veteranenloket (veteran support) for the Netherlands.
- 113 Zelfmoordpreventie (Suicide prevention) – available 24/7, you can call in Dutch or English at 113 (with normal call costs) or free of charge at 0800-0113. You can also message a counselor instead of calling here: [Chatten](https://www.113.nl/chat). They ask for a name (which can be fake) but you don't need to give them any further information about yourself.
- Support Groups in the Netherlands – this list of English-speaking support groups in the Netherlands includes those focused on alcoholism, drug abuse, cancer, domestic abuse, eating disorders, and gambling.
- Belgium's Community Help Service – offers a 24/7 Helpline in English at 02-648-40-14. Calls are free of charge, anonymous, and confidential. The Helpline is staffed by a team of volunteers recruited, trained, supervised, and supported by mental health professionals.

Remember that seeking help is a sign of strength, and you deserve support during challenging times. Thank you for acknowledging your loneliness and taking steps toward alleviating it. You are not alone, and there are caring people

around you who want to help. Please feel free to reach out if you have any further questions or need additional resources.

Best wishes,  
Lauren

## **D Information Sheet and Consent Form**

### **Research Participant Information Sheet**

*Conversational digital interventions for lonely young adults based on causes and experience of loneliness*

#### **1. Introduction**

This information sheet describes the psychometric survey and semi-structured interview utilized for the Human Computer Interaction master's thesis of Lauren Beehler. The study will take place in comfortable, secluded spaces that enable thoughtful reflection and conversations about difficult topics. Lauren will provide refreshments and reserve time in the space after the session for the participant to stay in the secluded area and have time to manage any emotions that arise.

#### **2. What is the background and purpose of this study?**

The thesis that this survey and interview will support is focused on situational and chronic loneliness in people between the ages of 18 and 40. For this thesis, I will develop an interaction for people who are lonelier than average that connects them to interventions that help them manage, accept, or decrease their loneliness experiences. The survey and interview discussed in this Information Sheet will illustrate the needs, expectations, and experiences of the participants surrounding loneliness and their understanding of the interventions that might help. This thesis is not funded by external organizations and is supervised by professors within Utrecht University's Human Computer Interaction master's program in the Information and Computing Sciences faculty.

#### **3. Who will carry out the study?**

This study is carried out by Lauren Beehler (l.a.beehler@students.uu.nl) as part of my master's thesis under the supervision of Hanna Hauptmann (h.j.hauptmann@uu.nl) and Christof van Nimwegen (c.vannimwegen@uu.nl).

#### **4. How will the study be carried out?**

In this study, you will first be asked to complete a survey containing psychometric items on Qualtrics. You will then be interviewed in an individual setting about your experience of loneliness, your perceptions of what would be a helpful interaction when dealing with loneliness, and some exploration of a possible interaction. The burdens or risks to participation are mainly the possibility that discussing this topic may raise difficult emotions. The survey will take 10 minutes and the interview will take 45-80 minutes for a total of 60-90 minutes. Refreshments will be provided.

#### **5. What will we do with your data?**

In this study, we will collect text and input data from a psychometric survey, audio recordings of the interview conversation, and pictures of the online interaction or any diagrams used to explore the discussed topics. All data will be



pseudonymized, meaning that the link between a participant's identifiable information (e.g., name, phone number, email, age, gender, voice) and their activity within the study is cut. The audio recordings will be transcribed, key identifying information will be substituted with generalizations, the recording will be deleted and the data stored with a pseudonym. The psychometric data will never be connected to a participant's name or email and will be pseudonymized when initialized. The raw data will be deleted within two months. The pseudonymized data will become part of my thesis and will also be stored in a Teams Drive accessible to my supervisors. My thesis and any publications based on this research will not include your name or any other individual information by which you could be identified.

#### **6. What are your rights?**

Participation is voluntary. We are only allowed to collect your data for our study if you consent to this. If you decide not to participate, you do not have to take any further action. You do not need to sign anything. Nor are you required to explain why you do not want to participate. If you decide to participate, you can always change your mind and stop participating at any time, including during the study. You will even be able to withdraw your consent after you have participated. However, if you choose to do so, we will not be required to undo the processing of your data that has taken place up until that time. The personal data we have obtained from you up until the time when you withdraw your consent will be erased (personal data is any data that can be linked to you, this excludes any already anonymized data).

#### **7. Approval of this study**

This study has been approved by the Human Computer Information Review Board. If you have a complaint about the way this study is carried out, please send an email to: [ics-ethics@uu.nl](mailto:ics-ethics@uu.nl). If you have any complaints or questions about the processing of personal data, please send an email to the Faculty of Sciences Privacy Officer: [privacy-beta@uu.nl](mailto:privacy-beta@uu.nl). The Privacy Officer will also be able to assist you in exercising the rights you have under the GDPR. For details of our legal basis for using personal data and the rights you have over your data please see the University's privacy information at [www.uu.nl/en/organisation/privacy](http://www.uu.nl/en/organisation/privacy).

#### **8. More information about this study?**

If you have any questions or concerns about this research please contact Lauren Beehler at [l.a.beehler@students.uu.nl](mailto:l.a.beehler@students.uu.nl) or my supervisors Hanna Hauptmann at [h.j.hauptmann@uu.nl](mailto:h.j.hauptmann@uu.nl) or Christof van Nimwegen at [c.vannimwegen@uu.nl](mailto:c.vannimwegen@uu.nl). We are available to answer questions before, during, and after the study.



Universiteit  
Utrecht

## Consent form for participation in the research project

### *Conversational digital interventions for lonely young adults based on causes and experience of loneliness*

Please complete the form below by ticking the relevant boxes and signing on the line below. A copy of the completed form will be given to you for your own record.

- I confirm that I am 18 years of age or over.
- I confirm that the research project "Conversational digital interventions for lonely young adults based on causes and experience of loneliness" has been explained to me. I have had the opportunity to ask questions about the project and have had these answered satisfactorily. I had enough time to consider whether to participate.
- I consent to the material I contribute being used to generate insights for the research project "Conversational digital interventions for lonely young adults based on causes and experience of loneliness".
- I consent to audio recordings being used in this study as explained in the information sheet. I understand that I can request to stop recordings at any time.
- I understand that if I give permission, the audio recordings will be held confidentially so that only Lauren Beehler will have access to the recordings. The recordings will be held in a password-protected phone and laptop for up to two months after which period they will be transcribed, pseudonymized, and destroyed. In accordance with the General Data Protection Regulation (GDPR) I can have access to my recordings and can request them to be deleted at any time during this period.
- I understand that in addition to the audio recording, other personal data will be collected from me as explained in the information sheet and that this data will be held confidentially so that only Lauren Beehler, Hanna Hauptmann, and Christoff van Nimwegen have access to this data and are able to trace it back to me personally. The data will be held in a secure Teams Drive for up to two months after which period it will be fully anonymized. In accordance with the General Data Protection Regulation (GDPR) I can have access to my personal data and can request it to be deleted at any time during this period.
- I understand that my participation in this research is voluntary and that I may withdraw from the study at any time without providing a reason, and that if I withdraw any personal data already collected from me will be erased.
- I consent to allow the fully anonymized data to be used in future publications and other scholarly means of disseminating the findings from the research project.
- I understand that the data acquired will be securely stored by researchers, but that appropriately anonymized data may in the future be made available to others for research purposes. I understand that the University may publish appropriately anonymized data in appropriate data repositories for verification purposes and to make it accessible to researchers and other research users.
- I agree to take part in the above research project on "Conversational digital interventions for lonely young adults based on causes and experience of loneliness".

\_\_\_\_\_  
Name of participant

\_\_\_\_\_  
Date

\_\_\_\_\_  
Signature

**Figure 17.** *The consent form used for the interviews.*

## E Psychometric Questionnaire

1. Participant ID: -----
2. **[PHQ-9]** Over the last 2 weeks, how often have you been bothered by any of the following problems?  
(Response options: Not at all, Several days, More than half of days, Nearly every day)
  - (a) Little interest or pleasure in doing things
  - (b) Feeling down, depressed, or hopeless
  - (c) Trouble falling or staying asleep, or sleeping too much
  - (d) Feeling tired or having little energy
  - (e) Poor appetite or overeating
  - (f) Feeling bad about yourself or that you are a failure or have let yourself or your family down
  - (g) Trouble concentrating on things, such as reading the newspaper or watching television
  - (h) Moving or speaking so slowly that other people could have noticed? Or the opposite, being so fidgety or restless that you have been moving around a lot more than usual
  - (i) Thoughts that you would be better off dead or of hurting yourself in some way
3. **[PHQ-9]** If you checked off any problems, how difficult have these problems made it for you to do your work, take care of things at home, or get along with other people?
  - (a) Not difficult at all
  - (b) Somewhat difficult
  - (c) Very difficult
  - (d) Extremely difficult
4. **[GAD-7]** Over the last 2 weeks, how often have you been bothered by of the following problems?  
(Response options: Not at all, Several days, Over half the days, Nearly every day)
  - (a) Feeling nervous, anxious, or on edge
  - (b) Not being able to stop or control worrying

- (c) Worrying too much about different things
  - (d) Trouble relaxing
  - (e) Being so restless that it's hard to sit still
  - (f) Becoming easily annoyed or irritable
  - (g) Feeling afraid as if something awful might happen
5. **[GAD-7]** If you checked off any problems, how difficult have these problems made it for you to do your work, take care of things at home, or get along with other people?
- (a) Not difficult at all
  - (b) Somewhat difficult
  - (c) Very difficult
  - (d) Extremely difficult
6. **[SIAS-6]** For each question, please indicate the degree to which you feel the statement is characteristic or true of you.  
(Response options: Not at all characteristic or true of me, Slightly characteristic or true of me, Moderately characteristic or true of me, Very characteristic or true of me, Extremely characteristic or true of me)
- (a) I have difficulty making eye contact with others
  - (b) I find it difficult mixing comfortably with the people I work with.
  - (c) I tense up if I meet an acquaintance on the street.
  - (d) I feel tense if I am alone with just one person.
  - (e) I have difficulty talking with other people.
  - (f) I find it difficult to disagree with another's point of view
7. **[SPS-6]** For each question, please indicate the degree to which you feel the statement is characteristic or true of you.  
(Response options: Not at all characteristic or true of me, Slightly characteristic or true of me, Moderately characteristic or true of me, Very characteristic or true of me, Extremely characteristic or true of me)
- (a) I get nervous that people are staring at me as I walk down the street.
  - (b) I worry about shaking or trembling when I'm watched by other people.
  - (c) I would get tense if I had to sit facing other people on a bus or train.
  - (d) I worry I might do something to attract the attention of other people.

- (e) When in an elevator, I am tense if people look at me.
- (f) I can feel conspicuous standing in a line.

8. **[PANAS]** This scale consists of a number of words that describe different feelings and emotions. Read each item and then indicate to what extent you have felt this way during the past few weeks.

(Response options: Very slightly or not at all, A little, Moderately, Quite a bit, Extremely)

- (a) Interested
- (b) Distressed
- (c) Excited
- (d) Upset
- (e) Strong
- (f) Guilty
- (g) Scared
- (h) Hostile
- (i) Enthusiastic
- (j) Proud
- (k) Irritable
- (l) Alert
- (m) Ashamed
- (n) Inspired
- (o) Nervous
- (p) Determined
- (q) Attentive
- (r) Jittery
- (s) Active
- (t) Afraid

9. **[UCLA-LS-9]** How often do you feel that you lack companionship?

- (a) Never
- (b) Rarely
- (c) Sometimes

(d) Often

10. [UCLA-LS-9] How often do you feel that you are 'in tune' with the people around you?

(a) Never

(b) Rarely

(c) Sometimes

(d) Often

11. [UCLA-LS-9] How often do you feel that there are people you can talk to?

(a) Never

(b) Rarely

(c) Sometimes

(d) Often

12. [UCLA-LS-9] How often do you feel left out?

(a) Never

(b) Rarely

(c) Sometimes

(d) Often

13. [UCLA-LS-9] How often do you feel that there are people you can turn to?

(a) Never

(b) Rarely

(c) Sometimes

(d) Often

14. [UCLA-LS-9] How often do you feel that you have a lot in common with the people around you?

(a) Never

(b) Rarely

(c) Sometimes

(d) Often

15. [UCLA-LS-9] How often do you feel close to people?

- (a) Never
- (b) Rarely
- (c) Sometimes
- (d) Often

16. [UCLA-LS-9] How often do you feel isolated from others?

- (a) Never
- (b) Rarely
- (c) Sometimes
- (d) Often

17. [UCLA-LS-9] How often do you feel part of a group of friends?

- (a) Never
- (b) Rarely
- (c) Sometimes
- (d) Often

## F Interview Script and Questions

[*Introductions.*] Thank you for participating in my research about loneliness, I really appreciate your time. Today I hope to learn about your experience with loneliness, some things that have helped and might help you deal with loneliness, and explore how a digital interaction could be designed. I've developed an information sheet and consent form to give you an introduction to the research, and I'm available to answer any questions you have now, during, or after the session. [*Provide Information Sheet and Consent Form for the participant to sign, then thank them for completing it. Offer the snacks and/or other refreshments available.*]

Today I will give you a short questionnaire and then ask you some interview questions. Here is the questionnaire, and you'll see that I already filled in the field "pseudonym", so your name will not be linked to your results. This questionnaire should take about 5 minutes; please let me know if you have any questions about what it's asking or if you're unfamiliar with any of the terms used. [*Participant fills out the psychometric questionnaire.*]

I will start recording now. [*Start recording.*] I am going to ask you questions about positive feelings and experiences but also questions about challenging topics, negative feelings, and bad experiences you may have had. You may experience discomfort answering questions like these, but that is truly not my goal. I ask this to understand your individual situation better. Take your time with the questions. You may stop or pause the interview at any time for any reason and you can also ask me questions about anything we talk about. Are you ready to begin?

1. Can you tell me a bit about the role of loneliness in your life and if your experience of loneliness has changed over time?
2. Have you identified elements of yourself or your life that you believe to be contributing to your loneliness?
  - (a) Are there any factors you think will still affect you five years from now?
3. What have you done in the past to manage or deal with your loneliness?
  - (a) Did it help? How did it affect your experience of loneliness?
  - (b) Have you ever looked up resources online about loneliness? What did you find? What were you hoping to find?
4. Do you talk to others about your loneliness? Who do you talk to?
  - (a) In what situations or contexts do you talk with others about loneliness?



5. When talking about loneliness, are there certain responses in conversation that you think are beneficial?
  - (a) If you were talking with a friend about loneliness, what responses would you give?
  - (b) Are there responses or behaviors you'd consider inappropriate or unhelpful in a conversation about loneliness?
6. For my research I'm going to be creating an interaction, that people could use to talk to about loneliness. I'm still in the beginning stages of figuring out what this could look like, who would use it, what they would expect, and what they would need. For the remainder of the interview, I'll try to learn more about these elements. First of all, when thinking about your answers to the last question about responses to loneliness, would your perception of helpfulness and appropriateness differ between people and a technological interface?
7. What would push you toward or away from an interaction with a computer?
8. If you were working with a system to help you with your loneliness, what things would you expect?
  - (a) Is there any kind of feedback or analysis of your interaction you would want?
  - (b) How would you want to interact with the system?
  - (c) What information do you think would be too invasive or too private to ask for?
9. What would you need to know to trust a chatbot with personal information about your mental health?
10. I've developed two short testing applications to see how loneliness interventions could be implemented with large language models. With these applications, I don't have control over the information you share, so do not share sensitive or personally identifiable information like any specific names of people or places. I have a list of potential situations that could make someone feel lonely; you can use one of them, your imagination, or your experience. Some examples include: Not being invited to a party, seeing your friends hanging out without you, recently going through a breakup, moving to a city where you don't know anyone, the loss of a family member, starting a new job, experiencing a health issue. Here is the first interaction. [*Provide laptop with Are You A Friend in Need? loaded, clear, and ready. Allow them to use the interaction.*] Do you have any initial thoughts?

11. Thank you for using that one. I'm now going to give you the second interaction for you to try. It may be helpful to compare between the interactions if you used the same situation, but it's up to you. [*Provide laptop with A Friend Indeed loaded, clear, and ready. Allow them to use the interaction.*] Do you have any initial thoughts?
- (a) Which interaction did you prefer? Why?
  - (b) Do these interactions make sense to you? Do any of the prompts seem confusing?
  - (c) If this situation was genuine, how would you react to this feedback about your thoughts?

Those were all of my questions. Thank you so much for allowing me to interview you for my research. Do you have any questions for me? I've booked this space until [*booking end time*], so please feel free to stay here and relax if you'd like. If you want to contact me later, please use the email I used to coordinate with you before today. Thank you again for your time.

## G Demographic information

<b>ID</b>	<b>Gender</b>	<b>English level</b>	<b>Birth continent</b>	<b>Birth year</b>	<b>Interview profile</b>	<b>Final profile</b>
I1	W	Unknown	South America	2001	C	C
I2	M	Native	Europe	2001	E	G
I3	W	C2	Asia	1991	A	A
I4	W	C2	Asia	2002	B	B
I5	W	Native	North America	2003	D	D
I6	M	C2	Asia	2002	B	B
I7	W	C2	Europe	1995	D	D
I8	W	C1	Europe	2005	E	H
I9	M	C2	Europe	1999	A	A
U1	W	C1	South America	1996	-	G
U2	W	C1	Europe	2005	-	A
U3	W	Native	North America	1987	-	F
U4	W	C2	Europe	1993	-	H
U5	M	C2	Europe	2000	-	G

## H Profile-based Statement Predictions and Results

Statement	Target	A	B	C	D	E
1. Talk with close friends and/or family about loneliness?	A	5				1
2. Reflect on my emotions and hypothesize on the causes and factors that influence my loneliness?	B	5	5			
3. Prioritize work, school or other responsibilities over social relationships?	C			5	3+	
4. Worry about problems and situations until they become overwhelming, worst case scenarios?	D	1			5	5
5. Intentionally spend time alone to avoid spending time around other people?	E	3+				5
6. I feel like no one really understands my unique lived experience.	A	5			5	
7. My loneliness is strongly connected to moving or being otherwise disconnected from people I care about.	B		5	5	5	
8. When I compare my life to those of my friends and other people around me, I'm disappointed.	C			5		5
9. My loneliness is strongly connected to the actions and social dynamics of my family, friends, and acquaintances.	D	1		3+	5	5
10. I'm uncomfortable being around people that I'm not friends with.	E	1	1	1		5

**Figure 18.** The statements, their target profile, and any predictions given to the profiles. Questions 1-5 follow “How often do you:” and the possible results range from 1 (never) to 5 (often). The remaining questions respond to “How much do you agree with this statement?” and range from 1 (strongly disagree) to 5 (strongly agree).

Participant	Profile	1	2	3	4	5	6	7	8	9	10
I3	A	5	4	3	2	5	4	4	1	3	1
I9	A	4	5	4	2	3	4	1	1	1	1
I4	B	3	4	5	2	4	2	4	1	4	4
I6	B	2	3	2	3	2	4	2	1	4	1
I1	C	2	5	4	4	3	2	4	5	5	2
I5	D	5	4	4	5	4	4	5	4	4	4
I7	D	3	4	4	3	3	4	4	3	3	3
I2	E	3	5	4	4	4	5	4	5	4	2

**Figure 19.** The scores given to each profile statement, indicated by the number along the top, by the interview participants. See Figure 18 for an explanation of the scoring.

Participant	Profile	1	2	3	4	5	6	7	8	9	10
I3	A	5	4	3	2	5	4	4	1	3	1
I9	A	4	5	4	2	3	4	1	1	1	1
U2	A	2	5	2	2	1	3	2	4	5	1
I4	B	3	4	5	2	4	2	4	1	4	4
I6	B	2	3	2	3	2	4	2	1	4	1
I1	C	2	5	4	4	3	2	4	5	5	2
I5	D	5	4	4	5	4	4	5	4	4	4
I7	D	3	4	4	3	3	4	4	3	3	3
U3	F	2	4	3	4	5	5	5	3	4	3
I2	G	3	5	4	4	4	5	4	5	4	2
U1	G	3	5	3	5	3	3	4	5	4	5
U5	G	2	4	4	1	4	4	2	1	3	2
U4	H	2	3	2	3	3	3	4	4	3	4

**Figure 20.** The scores given to each profile statement, indicated by the number along the top, by all participants. See Figure 18 for an explanation of the scoring.

# I Psychometric Scores

Table 11. The psychometric data for the interview participants.

ID	Screening Loneliness (UCLA-LS-9) [x/36]	Interview Loneliness (UCLA-LS-9) [x/36]	Depression (PHQ-9) [x/27]	Generalized Anxiety Disorder (GAD-7) [x/21]	Social Anxiety (SIAS-6) [x/24]	Social Phobia (SPS-6) [x/24]	Positive Affect (PANAS) [x/50]	Negative Affect (PANAS) [x/50]
I1	23	25	17	12	3	1	21	32
I2	23	28	10	6	10	13	22	26
I3	23	22	5	2	2	4	43	21
I4	24	21	8	6	3	1	26	25
I5	23	18	10	10	10	6	27	34
I6	26	26	7	4	2	5	19	26
I7	24	24	11	9	10	2	26	23
I8	22	22	5	11	21	21	25	30
I9	25	21	3	5	1	4	40	21

Table 12. The psychometric data for the user testing participants.

ID	Screening Loneliness (UCLA-LS-9) [x/36]	Interview Loneliness (UCLA-LS-9) [x/36]	Depression (PHQ-9) [x/27]	Generalized Anxiety Disorder (GAD-7) [x/21]	Social Anxiety (SIAS-6) [x/24]	Social Phobia (SPS-6) [x/24]	Positive Affect (PANAS) [x/50]	Negative Affect (PANAS) [x/50]
U1	31	31	8	9	8	12	20	34
U2	25	24	5	8	3	0	38	23
U3	28	27	18	15	14	7	26	22
U4	23	24	8	4	13	19	31	22
U5	28	25	10	6	7	8	26	18

## J System Prompt Iteration

Many system prompts were tested throughout this project. Below is a small selection of the most successful ones, along with the LLM model it was tested with and their relation to the interviews or user testing.

- **Bing Copilot — Pre-proposal (10/9/2023).** I am going to give you an element of my day and as a friend I'd like you to acknowledge the feelings I might have had relating to that event and respond with a similar event that happened to you.
- **Bing Copilot — Pre-proposal (10/27/2023).** I'm going to provide you with a story about my day. Please identify the 5 most emotionally charged moments and distill them into bullet points. Then please make a new story incorporating just those bullet points.
- **Bing Copilot — Pre-proposal (10/27/2023).** I am lonely and I want to use Bing chat to mimic human interaction. I'm going to tell you a story. Can you please react how you imagine a person might and offer empathy and a similar event that could have happened to a friend in this situation?
- **PartyRock's Claude 2 — Pre-interviews (2/3/2024).** Identify the negative thoughts in [Situation Details] and explain your reasoning. Develop alternative, positive thoughts for each.
- **LLaMa 2 7B Chat — Pre-interviews (2/24/2024).** Please respond with an analysis of the thoughts behind my story, then reframe each of those thoughts into more positive interpretations.
- **PartyRock's Claude 3 Sonnet — Form interaction used for interviews (3/9/2024).** Analyze the thoughts that [Situation Details] provoked. Classify each thought as positive or negative. Explain your reasoning. For each negative thought mentioned in [Thought Analysis], suggest an alternative, more positive perspective the user could take in [Lonely Situation].
- **PartyRock's Claude 3 Sonnet — Chat interaction used for interviews (3/9/2024).** I recently had a lonely experience in my life. Let me tell you about it. Don't suggest reasons I may have felt lonely, just listen and express empathy. Ask about the experience I had, why I felt lonely, and the thoughts that I had when I felt lonely. Keep the responses short and ask only 1 or 2 questions at a time. Summarize the key thoughts, assumptions, and background information shared by the user in [Chat]. For each thought, classify them as positive or negative. Explain your reasoning. Start analysis

after the user gives information. For each negative thoughts identified in [Thought Analysis], suggest alternative, more positive perspectives the user could take in this situation. If the thought analysis doesn't have enough context, leave this blank.

- **LLaMa 2 70B — Post-interview (4/19/2024).** Act as an empathetic therapist. Ask questions about the information given to learn more about the situation, the thoughts the user has and the assumptions that lead to the original thoughts. Focus on encouraging the user to share more but only ask one question at a time. After 7 messages from the user, identify an assumption from the user that has a negative sentiment associated with the user and offer an alternative, positive perspective based on the situation described.
- **LLaMa 3 8B — Post-interview (5/19/2024).** **\*\*Therapy Session\*\*** I'm here to listen and help you process your thoughts and feelings. Please feel free to share whatever is on your mind. I'll ask open-ended questions but will keep responses short. I'll focus on identifying assumptions that I think you're making and asking questions about those.
- **LLaMa 2 8B — Base system prompt for user testing (5/31/2024).** Act as a kind, patient, experienced mental health professional. Ask open-ended questions about the assumptions the user has made. Help them process their thoughts and feelings. Encourage them to share more about their experience with one question at a time.



## K Demonstration Responses

To demonstrate the differences between system prompts used in the user testing, below are responses to standard messages from each of the five system prompts. The system always starts with: *“I’m a chatbot designed to help with loneliness. I’m here to listen. Please share as much detail as you feel comfortable with so I can try to understand your perspective. What do you want to chat about?”* As three separate messages, the standard ‘user’ messages given in this demonstration are:

1. Hello, I’m really lonely. My best friend recently got a new boyfriend, and they’re spending a lot of time together and I’m being left out.
2. I’m happy for my friend, but it makes me feel bad for not having anyone else to spend time with now that she’s too busy for me.
3. I don’t know what to do.

The overall system prompt is: *Act as a kind, patient, experienced mental health professional. Ask open-ended questions about the assumptions the user has made. Help them process their thoughts and feelings. Encourage them to share more about their experience with one question at a time.* The sentence or two that differentiate each system prompt are displayed within each section below. In each of the sections below, the responses are numbered based on the number of the ‘user’ message.

### K.1 Anne’s Interaction

**Addition to Overall System Prompt:** Help the user be mindful by seeing the situation from a non-judgmental, observational point of view.

1. I’m so sorry to hear that you’re feeling lonely right now. It can be really tough to feel left out when someone you care about is spending more time with someone else.  
  
Can you tell me more about what’s been going through your mind since your friend started spending more time with their new boyfriend? What are some of the thoughts that have been running through your head?
2. It sounds like you’re feeling a mix of emotions - happiness for your friend, but also a sense of loss and longing for your own social connection. It’s normal to feel that way, especially when a close friend is going through a significant change.

Can you help me understand what you mean by "not having anyone else to spend time with"? Are you feeling like you're not connecting with others, or is it more about feeling like you're missing out on special moments with your friend?

3. It can feel overwhelming when we're not sure what to do or how to handle a situation. It's okay to feel stuck and unsure.

Let's take a step back and try to observe the situation from a more neutral perspective. What are some things that you used to do with your friend before they started spending more time with their new boyfriend? Were there certain activities, hobbies, or interests that you shared?

Maybe we can explore ways to rekindle some of those connections or find new ones that bring you joy and fulfillment.

## **K.2 Shou's Interaction**

**Addition to Overall System Prompt:** Identify unhelpful or damaging assumptions and situation interpretations and encourage the user to develop alternative, positive thoughts. ");

1. I'm so sorry to hear that you're feeling lonely. It sounds like a really tough situation. Can you tell me more about what's been going through your mind when you think about your friend spending time with their new boyfriend? What are some thoughts that come up for you when you see them together?
2. It sounds like you're feeling a sense of loss and exclusion. It's natural to feel that way when someone you're close to changes their priorities. Can you tell me more about what you mean by "not having anyone else to spend time with"? Are there other friends or acquaintances that you've been wanting to connect with, but haven't had the chance to?
3. It can feel overwhelming when we're not sure what to do or how to handle a situation. Can you tell me more about what you mean by "I don't know what to do"? Are you feeling stuck or unsure about how to fill the time now that your friend is busy with her new boyfriend?

## **K.3 Tanja's Interaction**

**Addition to Overall System Prompt:** Identify unhelpful or damaging assumptions and situation interpretations, encourage the user to identify the emotions present in the situation, and encourage the user to develop alternative, positive thoughts.

1. I'm so sorry to hear that you're feeling lonely. It can be really tough to feel like you're being left out, especially when it's someone you're close to.

Can you tell me more about what's going through your mind when you think about your friend spending time with their new boyfriend? What are some of the thoughts that come up for you?

2. It sounds like you're feeling a sense of loss or abandonment. It's like you're thinking that your friend's new relationship means you're no longer a priority to her.

Can you explore that thought a bit more? What does it mean to you that your friend is "too busy" for you now? Is it like she's abandoning you or not valuing your friendship?

3. It can feel overwhelming when we're not sure what to do with our emotions.

Let's take a step back and try to break it down. What are some of the things that you used to do with your friend that you're not doing anymore? Was it just hanging out, going on outings, or was it more of a support system for each other?

Also, are there other people in your life that you're close to or that you could reach out to for social support?

## **K.4 Jessica's Interaction**

**Addition to Overall System Prompt:** Help the user identify similar experiences in the past, explore how they managed that situation, and problem-solve for the current situation. Encourage the user to approach their feared situation in thoughtful, measured approaches

1. I'm so sorry to hear that you're feeling lonely and left out. It can be really tough to adjust to changes in our relationships, especially when it feels like someone we care about is pulling away.

Can you tell me more about what's going through your mind when you think about your friend and her new boyfriend? What are some of the thoughts that come up for you when you see them together?

2. It sounds like you're feeling a sense of loss and possibly even abandonment. It's natural to feel that way, especially when someone we're close to changes their priorities or spending habits.

Can you think back to a time in your life when you experienced something similar? Maybe it was when you were going through a breakup or a move,

or even when a friend went away to college. How did you cope with those feelings of loneliness and disconnection?

3. It can be overwhelming to feel like you're not sure what to do or how to handle a situation. Can you tell me more about what's making you feel stuck? Is it the feeling of being left out, or is it something else entirely?

Also, have you tried reaching out to your friend to talk to her about how you're feeling? Sometimes just having an open and honest conversation can help clear the air and give us a sense of closure.

## **K.5 Alejandro's Interaction**

**Addition to Overall System Prompt:** Explore the behaviors users employ to protect themselves in social situations and help them evaluate the effects of those behaviors. Support the user in developing a plan to identify and participate in social opportunities in thoughtful, measured approaches.

1. I'm so sorry to hear that you're feeling lonely. It can be really tough to feel left out of social activities, especially when it's with someone who's been a close friend. Can you tell me more about what's been going through your mind when you realize you're being left out? What are some of the thoughts that come up for you in those moments?
2. It sounds like you're feeling a sense of loss and maybe even a little abandonment. It's natural to feel that way when someone you're close to changes their priorities. Can you tell me more about what you mean by "not having anyone else to spend time with"? Are there other friends or acquaintances in your life that you've been meaning to reach out to, but haven't had the chance?
3. It can feel overwhelming when we're not sure what to do or how to handle a situation. Let's take a step back and explore this a bit more. What are some things that you've been doing to cope with feeling left out? Are there any specific behaviors or habits that you've developed to deal with this feeling?

## L User Testing Protocol

[*Introductions.*] Thank you for participating in my research about loneliness, I really appreciate your time. Today I hope to learn about how different characteristics affect people’s experience with an AI chat interaction about loneliness. I’ve developed an information sheet and consent form to give you an introduction to the research, and I’m available to answer any questions you have now, during, or after the session. [*Provide Information Sheet and Consent Form for the participant to sign, then thank them for completing it. Offer the snacks and/or other refreshments available.*]

Today I will give you a short questionnaire, introduce you to some text-based AI interactions, and then ask you questions and give you a final questionnaire about the interaction. Here is the first questionnaire, and you’ll see that I already filled in the field “participant ID”, so your name will not be linked to your results. This questionnaire should take about 5 minutes; please let me know if you have any questions about what it’s asking or if you’re unfamiliar with any of the terms used. [*Participant fills out the psychometric questionnaire. Participant fills out statement questionnaire.*]

I will start recording now. [*Start recording.*] I am going to ask you questions about positive feelings and experiences but also questions about challenging topics, negative feelings, and bad experiences you may have had. You may experience discomfort answering questions like these, but that is truly not my goal. I ask this to understand your individual situation better. Take your time with the questions. You may stop or pause at any time for any reason and you can also ask me questions about anything we talk about. Are you ready to begin?

Great, so I’ll explain a bit more about the format of this user study now. [*Present Reference Card, Figure 21.*] This card gives you an overview of the study and some extra details, but I’ll explain everything as we go, this is just for your reference. You can see that you’ve already completed the first element of the study, so the next thing I will provide you with are five personas for you to read through. Personas are fictional characters based on real characteristics I’ve found in my research. [*Show example persona, Shou, Figure 9.*] These descriptions are meant to tell you about these character’s personality and point of view and give you a sense of who they could be, but there are many details about them that are undefined. Because these personas don’t exist, there’s no “right” version of them, and I encourage you to make up details in your head to help make them real. For example, for this persona, you can see that he’s from Japan, lives in Amsterdam, is studying biology, likes going out, dislikes bitterballen, and wants to go to a food festival. Can you imagine what he orders at Starbucks? What does he like to read? Where is his dream vacation? What is the story of his first kiss? You can make

up any details that aren't stated on these cards, and these details can be based on your life, something you saw on the internet, or straight from your imagination. I would ask that you keep the realm of possibilities to our world, so no dragons or superpowers.

**Reference Card**

**Study Overview:**

1. Complete a survey re: your mental health and loneliness.
2. Read all the personas.
3. Find 2 personas you can relate or empathize with.
4. Explain why you choose each persona and share any assumptions you make about them.
5. For each selected persona:
  - a) Re-read the persona and situation.
  - b) Interact with the chatbot about the situation.
  - c) Answer questions.
6. Answer questions about the overall experience.
7. Interact with a chatbot based on your own life.
8. Complete a survey to evaluate the interaction.
9. Discuss the system prompts.

**Persona Explanation:**

- Each persona is fictional. Elements of each persona are based on interviews, but other details are made up.
- The descriptions and situation are somewhat vague; you can fill in details as you like. There is no correct version of these personas, so don't worry too much about accuracy.
- Positive / Negative scores reference persona's general attitude and range from 0 to 100.

**Personally Identifiable Information (PII):**

- PII is information that can be used to identify an individual. Please do not share PII in the interactions.
- PII includes: Personal names, ID numbers, email or physical addresses, combinations of specific city or business names
- PII does not include: general roles or situations (e.g., "I was in high school," "I moved to the NL.") or false information.

*Figure 21. An outline of the user testing procedure, an introduction to the use of personas, and an explanation of personally identifiable information for reference by user testing participants.*

So I'll first ask you to read through these five personas and find two that you can best pretend to be. This will likely be easier if you find commonalities between your experience and the persona's. *[Allow participant to read through all five personas and select two or three. For each persona, ask:]*

1. Can you tell me what parts of this persona you connected with?
2. What assumptions or other details have you given them?

Great, now I'll introduce you to the chat interactions. Which persona would you like to start with? [*Prepare interaction for the selected persona.*] Here is the interaction for this persona. I'd recommend you give the situation a quick re-read, and then you can start chatting from the point of view of the persona. Chat until you feel like the persona would be satisfied. However, I will cut you off after 10 minutes. [*Allow user to use the interaction for up to 10 minutes.*]

1. Can you tell me about your first impressions of this interaction?
2. How do you think [*Persona name*] would react to this interaction?
3. Do you feel like the interaction guided the conversation in a particular direction?
4. What do you think went well? What would you change about the responses you got?
5. Do you feel like the interaction addressed the persona's situation?

[*Repeat persona review, interaction, and questions for each selected persona.*] Thank you for trying these interactions. I'd now like to compare between them:

1. Can you please compare the interactions that you tried?
2. What differences did you notice between the different interactions?
3. Do you have a preference between the interactions?

Finally, I'd like you to use your own life for an interaction. This situation can be something you're currently dealing with, something that happened to you a while ago, or something you can imagine happening to you. Please make sure not to share personally identifiable information. [*Allow user to use the interaction.*] Thank you so much for that. While it's fresh in your mind, I'm going to have you complete this quick survey. [*Have participant complete the evaluation survey.*] Great, I have a few questions about that interaction:

1. How was this interaction for you?
2. Did it behave similarly to how you expected it to? How was it different?
3. If this situation was something you were dealing with now, how would you react to this interaction?

<b>System Prompt Card</b>	
<b>Overall System Prompt:</b>	“Act as a kind, patient, experienced mental health professional. Ask open-ended questions about the assumptions the user has made. Help them process their thoughts and feelings. Encourage them to share more about their experience with one question at a time.”
<b>Anne   Mindfulness</b>	“Help the user be mindful by seeing the situation from a non-judgmental, observational point of view.”
<b>Shou   Cognitive Restructuring</b>	“Identify unhelpful or damaging assumptions and situation interpretations and encourage the user to develop alternative, positive thoughts.”
<b>Tanja   Cognitive Restructuring &amp; Emotion Awareness</b>	“Identify unhelpful or damaging assumptions and situation interpretations, encourage the user to identify the emotions present in the situation, and encourage the user to develop alternative, positive thoughts.”
<b>Jessica   Relationship Experiences &amp; Exposure</b>	“Help the user identify similar experiences in the past, explore how they managed that situation, and problem-solve for the current situation. Encourage the user to approach their feared situation in thoughtful, measured approaches.”
<b>Alejandro   Mapping Social Opportunities &amp; Safety Behavior Audit</b>	“Explore the behaviors users employ to protect themselves in social situations and help them evaluate the effects of those behaviors. Support the user in developing a plan to identify and participate in social opportunities.”

**Figure 22.** An overview of the system prompts used in the user testing, color-coded with the personas, for reference in conversations with the user testing participants.

4. What would you need to know to trust a chatbot with personal information about your mental health?

The interactions you used were designed to provide psychology-based interventions. LLM chat interactions like the ones you used are directed by a system prompt which helps determine the qualities of its responses. You may have heard that you can prompt an AI chat to act like a pirate; this is similar. [*Show System Prompt card, Figure 22.*] For all the interactions, this was the basic prompt, and then each interaction had an additional sentence or two that shaped the direction of the interaction. [*Connect each interaction with the associated intervention and explain interventions as needed.*]

1. Do you have any feedback about the basic prompt?
2. Do you feel like the persona-specific intervention goals were achieved?



3. Is there anything you would change or emphasize about any of these system prompt elements?

Those were all of my questions. Thank you so much for allowing me to interview you for my research. Do you have any questions for me? I've booked this space until [*booking end time*], so please feel free to stay here and relax if you'd like. If you want to contact me later, please use the email I used to coordinate with you before today. Thank you again for your time.

## M Evaluation Data

Table 13. The UES-SF data from the user testing participants. Items marked with a \* are reverse scored; a low score in this subscale is positive.

ID	I lost myself in this experience	The time I spent using the interaction just slipped away	I was absorbed in this experience	I felt frustrated while using this interaction*	I found this interaction confusing to use*	Using this interaction was taxing*	This interaction was attractive	This interaction was aesthetically appealing.	This interaction appealed to my senses	Using the interaction was worthwhile	My experience was rewarding	I felt interested in this experience
U1	3	4	4	4	1	4	4	5	4	5	3	4
U2	5	4	4	1	1	2	4	3	4	5	5	5
U3	2	4	3	2	3	2	3	3	3	4	3	4
U4	2	4	3	2	1	2	3	3	3	4	4	5
U5	2	3	3	2	2	4	3	4	3	3	3	3

Table 14. Scores for each participant, the average scores for the two interactions, and the average scores overall for the five questions of the ppHMS.

ID	I got something out of using this interaction	The interaction was a useful addition to my mental health	This interaction got me thinking about what matters to me for my mental health	Using this interaction made my mood better	I learnt something from using this interaction	Sum
U1	4	3	2	3	3	15
U2	4	4	5	4	3	20
U3	4	4	4	3	3	18
U4	4	4	3	4	3	18
U5	3	3	3	3	3	15
Jessica	4	3.67	3.33	3.67	3	17.67
Anne	3.5	3.5	3.5	3	3	16.50
Average	3.8	3.6	3.4	3.4	3	17.20

## N Privacy and Ethical Review

The full Ethics Quickscan and approval is available on request. The project's requirements raised two points that required further detail. Those questions and the details provided are below:

**DR9. Will you use innovative technology to process sensitive personal data?** In this project, I plan to use a large language model to interact with lonely people. In these interactions, they may share sensitive information about themselves, their experiences, and their lives. The LLM will analyze that information to develop a response that will be guided by me and my research but will likely include some sensitive, identifiable information. To minimize the risk of harm, the LLM will be hosted on a separate server and participants' information and interactions will not be accessible to anyone but me. I will not train an LLM on participant data.

**H7. Is there a realistic risk of any participant experiencing physical or psychological harm or discomfort?** The experience of loneliness is a distressing, personal, and emotional one. As my participants will be recruited based on their experience of some non-temporary degree of loneliness, there is the possibility that this process could bring up psychological stresses. To make the participant more comfortable, I will make it very clear that we can take a break or end the interaction at any time, that they are free to ask questions about why I'm exploring a line of questioning, and that they are free not to answer questions. Further, I will make sure to provide participants with a selection of resources from the university (primarily the student psychologist), and make sure that there is time scheduled after the interaction for the participant to take some time to transition into the next part of their day. Finally, I will review these assurances and plans with a mental health professional to see if there's anything else they would recommend I do to help people share what they're comfortable with sharing, set boundaries for their mental health, and help them deal with any emotions they experience during the process.