A Mobile Intervention to Reduce Anxiety Among University Students, Faculty, and Staff: Mixed Methods Study on Users’ Experiences

Sarah Livermon¹*, Audrey Michel²¶, Yiyang Zhang²¶, Kaitlyn Petz², Emma Toner², Mark Rucker¹, Mehdi Boukhechba¹, Laura E. Barnes¹, and Bethany A. Teachman²*

¹Department of Systems and Information Engineering, University of Virginia, Charlottesville, United States of America
²Department of Psychology, University of Virginia, Charlottesville, Virginia, United States of America

* Corresponding author
E-mail: sat2ew@virginia.edu (SL)
E-mail: bteachman@virginia.edu (BAT)

†These authors contributed equally to this work.
Abstract

Anxiety is highly prevalent among college communities, with significant numbers of students, faculty, and staff experiencing severe anxiety symptoms. Despite this high prevalence, many individuals remain untreated due to barriers such as time, stigma, waitlists, and cost of services. Digital mental health interventions (DMHIs), including Cognitive Bias Modification for Interpretation (CBM-I), offer promising solutions to enhance access to mental health care, yet there is a critical need to evaluate user experience and acceptability of DMHIs. This study used a mixed methods approach to gather feedback from users of the first trial of a mobile application called "Hoos Think Calmly" (HTC), which offers brief CBM-I training to reduce negative interpretations and increase flexible thinking in response to stressors commonly experienced by students, faculty, and staff at a large public university. Following the parent trial (https://osf.io/36grh/) qualitative data was collected through semi-structured interviews from a subset of participants (n=22). Thematic analysis revealed five main themes: Effectiveness of the Training Program; Feedback on Training Sessions; Barriers to Using the App; Use Patterns; and Suggestions for Improvement. Feedback highlighted the importance of greater content relatability and personalization, while also identifying forgetfulness and not understanding the intervention format or rationale as barriers to using the program. Participants tended to use the program at routine or scheduled times rather than during specific moments of stress or anxiety and relied heavily on the app’s notification system. Suggestions for improvement focused on incorporating progress tracking, offering greater customization options, and integrating more diverse training content. Additionally, biweekly user experience questionnaires sent to all participants in the active treatment condition (n=134) during the parent trial showed most participants reported the program to be slightly to somewhat helpful in reducing or managing their anxiety or stress. Findings highlight the importance of understanding users’ experience and iterative DMHI development.
Author Summary

Effective digital mental health interventions can help address the high levels of reported anxiety in college communities. These interventions can be used to reduce the barriers to accessing mental health services and promote flexible thinking to reduce the harmful impacts of anxiety. Despite evidence supporting the effectiveness of various DMHIs, their impact is often hindered by high attrition rates alongside low engagement. In this study, we gathered mixed-methods feedback data on a new DMHI called Hoos Think Calmly to assess individual user experiences that may impact engagement. Our goal is to better understand participants’ evaluation of HTC and to gather feedback on how to enhance acceptability of the program in future iterations. Data was collected through feedback interviews and user experience questionnaires. We found that participants perceived HTC to be somewhat effective at reducing or managing their anxiety, but also noted barriers to ongoing use. Our study provides helpful insights into app use patterns, barriers to engagement, and general suggestions for improving DMHIs.

Introduction

Background

Anxiety disorders are highly prevalent among college communities. A 2023 study indicated that 41.8% of college students and 38.6% of faculty and staff had experienced symptoms consistent with severe or extremely severe anxiety over the past week [1]. Despite these alarmingly high rates, many individuals in need of mental health services remain untreated [2]. Some of the most common barriers to receiving mental health services among college students are access, time, and stigma. Accessibility barriers specifically include lengthy wait times, the cost of services, and perceived lack of options [3,4]. Similarly, university faculty reported stigma and lack of support as their main barriers to receiving mental health care [5]. In light of these barriers, there has been increased interest in digital mental health interventions (DMHIs). DMHIs provide mental health support via technologies like web-based or mobile
platforms and have the potential to expand treatment delivery options and enhance access to mental health care [6]. Studies show DMHIs may be particularly effective at reducing depression, anxiety, and stress outcomes in academic settings [7,8].

The use of Cognitive Bias Modification for Interpretation (CBM-I) is a mechanistic intervention that may be an effective component of treatment for anxiety [9-11], given extensive research showing that selective threat interpretations are common among individuals high in anxiety [12]. CBM-I is a type of training designed to modify threat-focused cognitive biases related to interpretation processes (i.e., assigning threatening meanings to ambiguous situations). CBM-I training typically involves presenting participants with ambiguous scenarios or stimuli and encouraging them to interpret them in a more positive or benign manner. Through repeated practice and reinforcement, the aim is to help individuals develop more flexible patterns of interpretation that are not so rigidly negative, which is expected to then reduce anxiety.

Although there is evidence supporting the effectiveness of various DMHIs, such as online CBM-I training (including at our team’s MindTrails website; [13-15]), their impact is hindered by high attrition rates alongside low engagement [16-18]. Better ways to overcome these common limitations of DMHIs need to be better understood. A 2021 review found that participants’ engagement with DMHIs increased when users found the content appealing, perceived the program to be a good fit, and experienced perceived changes in their mental health [19]. Thus, we need to understand what makes content appealing and a good fit to users. Research on DMHIs tends to rely on quantitative metrics to evaluate outcomes, which are very valuable but limited in their ability to contextualize data and assess individual user experiences that may impact engagement. This points to the value of also using qualitative approaches to help design, evaluate and improve DMHIs [20-22]. For instance, qualitative methods can be used during the post-trial phase of pilot studies to gain insight into users’ acceptance of and feedback on DMHIs [23]. This feedback can be especially valuable to guide improvements for future versions of a DMHI.
Objective
The objective of this study is to collect feedback from participants in the first trial of a mobile application called “Hoos Think Calmly” (HTC) to evaluate acceptability and gain a deeper understanding of their experiences and suggestions for improvement. HTC is designed to increase flexible thinking in response to stressors commonly experienced by students, faculty, and staff at a large public university through brief ~5 minute interventions (microdoses) of CBM-I training delivered by mobile phone to university community members. Study design and analysis plans were preregistered on Open Science Framework [24].

Methods
Ethics approval and consent to participate
This study was approved by the Institutional Review Board for Social and Behavioral Sciences (IRB-SBS) at the University of Virginia. All participants provided written informed consent via an electronic consent form prior to enrollment and participation.

Strategies to understand and mitigate bias: Reflexivity and credibility
Our team acknowledges that we bring biases in favor of HTC, having grants that fund the program and prior tests of the parent program MindTrails that demonstrated efficacy in reducing anxiety [13-15]. It is likely that these factors may predispose us towards interpreting the current findings more positively.

To bolster methodological integrity, we employed several credibility strategies. Two of the initial coders on the research team were individuals not involved in the data collection process to help mitigate biases and allow for fresh perspectives [25]. Likewise, we established an intercoder reliability score exceeding 0.80 on a subset of interviews before proceeding with general coding to increase coding consistency. Throughout the analysis, the lead author attended a recurring journal club focused on best practices in qualitative research. During this time, the lead author also spent time examining personal judgements and beliefs to better understand how biases
might influence the research process and outcomes. We encouraged peer examination, and colleagues were asked to give feedback on our research methodology and procedures. An audit trail was kept of all edits made to the codebook throughout the analysis process to help establish confirmability and trustworthiness of data [see 26]. This study was reported using the Standards for Reporting Qualitative Research (SRQR) guidelines [27] (S1 Text).

**Intervention**

In the parent randomized control trial (RCT), n=134 participants assigned to the active treatment condition were notified to complete two microdose sessions per day over a 6-week period. Prior to beginning a microdose session, participants completed a pre-intervention survey in which they reported their current mood (and completed a very brief imaginal exercise depending on their anxiety level) and selected a specific stressor domain for targeted training. Domain options included academics/work/career development, social situations, romantic relationships, finances, physical health, mental health, family and home life, and discrimination. The content of the training scenarios was modified to match common situations encountered for the different groups (e.g., to capture fears of negative evaluation, the undergraduate group would read about giving a class presentation while the staff would read about presenting at a staff meeting). Post-intervention surveys were also administered to assess changes in mood and provide additional resources and tips for emotion regulation and application of the CBM-I learning to daily life. At the end of the day, participants were instructed to complete a nightly survey inquiring about their mood and ability to think flexibly throughout the day. During this survey, participants had the option to pre-select notification times for the next day’s microdose sessions.

A microdose session consisted of ~ 5 minutes of CBM-I training that aimed to reduce rigidly negative thinking patterns associated with anxiety. Training scenarios could consist of 10 CBM-I short scenarios (this occurred in the majority of sessions), or 1 long scenario, or 1 write-your-own scenario (except for the Discrimination domain, which did not include CBM-I; see 26). The variation was included to reduce repetitiveness and to introduce different tasks to increase
flexible thinking and the personal relevance of training. A short scenario was typically three sentences in length and consisted of an emotionally ambiguous story that raised a potential threat but was then resolved when the user completed the final word in the last sentence. The emotional resolution was reinforced through a subsequent comprehension question (see Figure 1). A long scenario was the same as a short scenario except it also included a range of thoughts, feelings, or behaviors that could occur in a particular situation to help with seeing multiple different perspectives. Users were then prompted to write down their own thoughts that could be helpful in that situation. The write-your-own scenario invited participants to craft their own short CBM-I scenario based on a personally relevant anxiety-inducing situation. Users were then tasked with resolving the scenario in a positive or non-threatening way and coming up with reasons for why the chosen resolutions could be likely to occur. Figure 1 displays a screenshot of the app’s homepage, domain selection page, and a short scenario example. More details on the intervention are available via Open Science Framework [28, 29].

Fig 1. Screenshots of the HTC app

![screenshots of the HTC app](https://doi.org/10.1101/2024.08.05.24311492)
Study Procedures and Data Collection

This study used quantitative and qualitative approaches to understand user experience and acceptability of the HTC mobile health (mHealth) platform. Data was collected via questionnaires and semi-structured interviews. Four biweekly user experience questionnaires were administered to all participants assigned to the treatment condition throughout their study participation. The first three questionnaires were administered every two weeks during the 6-week study period, and the final questionnaire was administered at the 8-week mark as a post-study follow-up. In addition to survey questionnaires, a subset of participants who took part in the 6-week pilot RCT to test the HTC mobile application were selected to participate in a user feedback interview. The selection process involved purposive sampling based on group assignment, study progress, and time passed since enrollment. Participants who were assigned to the treatment condition and had completed at least 3 microdose sessions were randomly selected to participate in a feedback interview. The interview selection process was initiated by inviting participants with the closest study enrollment date to the time of recruitment and moving backwards until our target accrual number was met. We aimed to complete approximately 20 interviews based on norms in the literature for papers with similar goals [30, 20] and also due to practical considerations. A total of 89 interview invites were sent to faculty members, staff members, graduate students, and undergraduate students, and 22 interviews were completed using a semi-structured interview guide (S2 Text). Participant characteristics are provided in Table 1.
### Table 1. Characteristics of interviewee participants and survey respondents

<table>
<thead>
<tr>
<th>Variable</th>
<th>Questionnaire sample (n=95)</th>
<th>Interview sample (n=22)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>%</td>
</tr>
<tr>
<td>Role at university</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Undergraduate</td>
<td>26</td>
<td>27.4</td>
</tr>
<tr>
<td>Graduate</td>
<td>31</td>
<td>32.6</td>
</tr>
<tr>
<td>Faculty</td>
<td>6</td>
<td>6.3</td>
</tr>
<tr>
<td>Staff</td>
<td>32</td>
<td>33.7</td>
</tr>
<tr>
<td>Gender/Sex (the item did not distinguish between gender identity and sex at birth)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Woman</td>
<td>76</td>
<td>80.0</td>
</tr>
<tr>
<td>Man</td>
<td>12</td>
<td>12.6</td>
</tr>
<tr>
<td>Transgender Man</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Transgender Woman</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Other identity</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>No answer provided</td>
<td>7</td>
<td>7.4</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White/European origin</td>
<td>63</td>
<td>66.3</td>
</tr>
<tr>
<td>South Asian</td>
<td>7</td>
<td>7.4</td>
</tr>
<tr>
<td>East Asian</td>
<td>7</td>
<td>7.4</td>
</tr>
<tr>
<td>Black/African origin</td>
<td>4</td>
<td>4.2</td>
</tr>
<tr>
<td>More than one race</td>
<td>1</td>
<td>1.1</td>
</tr>
<tr>
<td>Other or Unknown</td>
<td>6</td>
<td>6.3</td>
</tr>
<tr>
<td>No answer provided</td>
<td>7</td>
<td>7.4</td>
</tr>
<tr>
<td>Ethnicity</td>
<td>83</td>
<td>87.4</td>
</tr>
<tr>
<td>-----------------------------------</td>
<td>----</td>
<td>------</td>
</tr>
<tr>
<td>Not Hispanic or Latino</td>
<td>5</td>
<td>5.3</td>
</tr>
<tr>
<td>Hispanic or Latino</td>
<td>7</td>
<td>7.4</td>
</tr>
<tr>
<td>No answer provided</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Measures at study baseline**

<table>
<thead>
<tr>
<th>Measure</th>
<th>Mean (SD)</th>
<th>Mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>OASIS</td>
<td>6.39 (2.87)</td>
<td>6.52 (2.42)</td>
</tr>
<tr>
<td>PHQ-2</td>
<td>1.63 (1.46)</td>
<td>1.71 (2.42)</td>
</tr>
<tr>
<td>PSS-4</td>
<td>6.88 (3.10)</td>
<td>7.48 (3.39)</td>
</tr>
</tbody>
</table>

OASIS = Overall Anxiety Severity and Impairment Scale; Scores range from 0-20. A higher OASIS score indicates more severe anxiety symptoms.

PHQ-2 = Patient Health Questionnaire-2; Scores range from 0-6. A higher PHQ-2 score indicates more severe depression symptoms.

PSS-4 = Perceived Stress Scale-4; Scores range from 0-16. A higher PSS-4 score indicates more stress.

SD = standard deviation

**Analysis**

Following our pre-registration [21], interview data was analyzed using thematic analysis as outlined by Braun and Clarke [31]. All interviews were transcribed using Descript audio transcription service [32]. Each transcript was then checked against the original audio recording by a member of our research team. This was for the purpose of ensuring transcription accuracy and enhancing team familiarity with the data. During transcript review, team members noted initial themes found in the data.

After all interview transcripts had been reviewed, preliminary codes were discussed, and an initial codebook was created. We then used an iterative coding process, allowing for continual addition or deletion of codes to align with our data. To most effectively capture context and meaning, text was coded at the “level of meaning”, allowing a code to encapsulate a line, sentence, or paragraph [33]. Interviews were coded using Delve qualitative data analysis software [34]. Four coders were trained by conducting inter-coder reliability (ICR) tests. Inter-
Coder reliability was measured using Krippendorff’s alpha, a measure of reliability in content analysis [35]. We underwent six rounds of ICR testing and reached an ICR rating of 0.83. Each round of testing involved independent coding of the same transcript by each coder, followed by comparison and discussion of coding differences. Transcripts from different participant groups (faculty, staff, graduate, and undergraduate) were selected for each round of ICR testing to ensure representation of the entire data set [36]. Changes were made to the codebook throughout this process. Each codebook iteration, including the final version, can be found on Open Science Framework [37]. After an inter-coder reliability score greater than 0.80 had been established, official coding of the transcripts commenced. Due to prolonged ICR training, only two of the four trained coders were available for the official coding process. Given time constraints and the high ICR rating, deviations were made to the preregistration, with each transcript being coded by one of the two remaining coders instead of undergoing double coding.

Weekly meetings were held to address coding challenges and refine the codebook. The lead author served as a moderator to help resolve disputes or provide clarification on coding issues. Thematic saturation was attained after analyzing 15 transcripts and was determined when no additional codes could be added to the codebook [38]. The final codebook consisted of 47 main codes (not including sub-codes) that were grouped into 10 themes. An initial thematic map was created to visualize the codes and their relationships. Of those 46 codes, 13 appeared in 50% or more interviews. A revised thematic map was created using the 13 most saturated codes, and 5 main themes were assigned (S1 Fig.). The full list of codes and definitions can be found in the codebook. Survey data was analyzed using descriptive statistics, including frequency counts and percentages.

**Results**

The results are categorized based on the 5 themes found in Table 2, along with examples of deidentified illustrative quotes. Considering the primary research goal for this paper was...
learning users’ impressions and suggestions from the qualitative feedback, and the survey data was included for triangulation purposes, we decided prior to the quantitative analyses to only examine those survey questions that aligned with the qualitative themes. Survey data are reported in Table 3.

Table 2. Main Themes Identified and Frequency of Codes among Interview Participants

<table>
<thead>
<tr>
<th>Theme</th>
<th>Code</th>
<th>Total (n=22)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effectiveness of the Training Program</td>
<td>Application of new coping strategies</td>
<td>11 (50%)</td>
</tr>
<tr>
<td></td>
<td>Perceived changes in stress/anxiety levels</td>
<td>14 (67%)</td>
</tr>
<tr>
<td>Feedback on Training Sessions</td>
<td>Relatability</td>
<td>21 (95%)</td>
</tr>
<tr>
<td></td>
<td>Repetitiveness</td>
<td>12 (54%)</td>
</tr>
<tr>
<td></td>
<td>Length</td>
<td>18 (82%)</td>
</tr>
<tr>
<td></td>
<td>Write-your-own</td>
<td>14 (64%)</td>
</tr>
<tr>
<td>Barriers to Using the App</td>
<td>Lack of understanding of intervention format or rationale</td>
<td>12 (55%)</td>
</tr>
<tr>
<td></td>
<td>Forgetfulness</td>
<td>12 (55%)</td>
</tr>
<tr>
<td>Use Patterns</td>
<td>Used routinely</td>
<td>15 (68%)</td>
</tr>
<tr>
<td></td>
<td>Used when reminded</td>
<td>11 (50%)</td>
</tr>
<tr>
<td></td>
<td>Chosen academics/work/career development domain</td>
<td>18 (82%)</td>
</tr>
<tr>
<td>Suggestions for Improvement</td>
<td>Feature Suggestions</td>
<td>19 (86%)</td>
</tr>
<tr>
<td></td>
<td>Domain Suggestions</td>
<td>16 (73%)</td>
</tr>
</tbody>
</table>
Table 3: Results from Questionnaires

<table>
<thead>
<tr>
<th>Survey question</th>
<th>1 = Not at all</th>
<th>2 = Slightly</th>
<th>3 = Somewhat</th>
<th>4 = Mostly</th>
<th>5 = Very</th>
<th>Mean ± SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. How helpful do you find Hoos Think Calmly for reducing your anxiety?</td>
<td>35 (15.63%)</td>
<td>57 (24.45%)</td>
<td>97 (43.30%)</td>
<td>30 (13.39%)</td>
<td>5 (2.23%)</td>
<td>2.61 ± 0.98</td>
</tr>
<tr>
<td>2. How much could you relate to the stories presented during training sessions?</td>
<td>9 (4.31%)</td>
<td>57 (27.28%)</td>
<td>76 (36.36%)</td>
<td>46 (22.01%)</td>
<td>19 (9.09%)</td>
<td>3.04 ± 1.02</td>
</tr>
<tr>
<td>3. How believable did you find the stories’ endings in terms of those outcomes happening in your life?</td>
<td>8 (3.83%)</td>
<td>52 (24.88%)</td>
<td>67 (32.06%)</td>
<td>60 (28.71%)</td>
<td>18 (8.61%)</td>
<td>3.14 ± 1.02</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Survey question</th>
<th>1 = Way too short</th>
<th>2 = Somewhat too short</th>
<th>3 = The right length</th>
<th>4 = Somewhat too long</th>
<th>5 = Way too long</th>
<th>Mean ± SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>4. I find the [length of the] training sessions to be:</td>
<td>0 (0.00%)</td>
<td>1 (1.09%)</td>
<td>46 (50.00%)</td>
<td>41 (44.57%)</td>
<td>2 (2.17%)</td>
<td>3.49 ± 0.57</td>
</tr>
</tbody>
</table>

*Percentages do not add up to 100 when “Prefer not to answer” was endorsed

Theme 1: Effectiveness of Training Program

Application of New Coping Strategies

Of the interview participants, 50% (11/22) expressed an ability to effectively apply new coping strategies (e.g. flexible thinking) that they learned from the program to their real-life situations.

There was this scenario that is about the advisor frowning in front of you and that instinctively made me think that I did something wrong or what I said is something
stupid, but the answer was that people frown when they think about something, and that scenario really helped me to think differently when I present my work in front of other people or in front of advisors. [Grad 13]

Perceived Changes in Stress or Anxiety

In addition, 64% (14/22) of interview participants discussed the program's impact on their stress or anxiety. The majority of those participants (8/14, 62%) expressed an overall positive change in symptoms, stating a reduction in their stress or anxiety levels.

It [the program] definitely did help…I immediately go to the worst case scenario, so I thought it helped like in the heat of the moment for me to redirect my train of thought to somewhere that’s more productive and less doomsday. [Undergrad 35]

Some of those participants (5/14, 38%) reported no change to their stress or anxiety while using the app.

This just didn't do anything for me. It just added another thing in my day that I had to think about and take care of that didn't feel like it was giving me anything back. [Undergrad 39]

Additionally, two participants (2/14, 14%) expressed a negative change in symptoms, reporting an increase in their stress or anxiety levels from use of the app.

Some [scenarios] seemed to cause me to worry more about things. Like I understand that it’s like guiding you through stressful situations and it’s not always gonna have a positive outcome for everything, but some of them were a little bit like, Ooh, that’s kind of scary to think about. [Undergrad 28]

It is worth noting that one participant reported experiencing both positive and negative changes in their stress or anxiety levels. Therefore, there are a total of 15 responses indicating positive, negative, or no change, out of the 14 participants who commented on the program's impact.

Additionally, a survey questionnaire was sent to all participants at the 2nd, 4th, 6th, and 8th week marks of their study participation to assess the effectiveness of the HTC program in
reducing their anxiety. A total of 225 responses were received (participants submitted
responses on multiple occasions, hence the greater number of responses than participants),
with the most common response (97/225, 43%) being that HTC was somewhat helpful at
reducing their anxiety (see Table 3).

Theme 2: Feedback on Training Sessions

Relatability

Out of the 22 interview participants, 95% (21/22) discussed the relatability of the training
scenarios. Among those 21 participants, approximately half (11/21, 52%) noted that the training
scenarios resonated with their own experiences and feelings.

I thought they were really [applicable], and sometimes I would even laugh because they
would be so accurate that I'm like, yeah, I went through that like last week or so. [Grad
49]

And I thought especially for academics, the situation or the stories that were there were
quite relatable. Like I thought, yeah, this is exactly what I'm feeling right now or what I'm
stressed about right now. [Grad 35]

Two thirds of participants (14/21, 67%) indicated that they sometimes did not find the training
scenarios relatable.

I felt like the demographic was younger than I am. It was a lot of you've got a new baby,
you've got a young child, or you've got parents who disapprove of what you're doing. My
parents are dead. This is not a thing I'm thinking about now… It just tended to not be
applicable to me. [Faculty 8]

There were times where I thought, oh, this doesn't relate to me at all, and why am I
taking time to do this? [Staff 39]
Notably, 4 participants mentioned finding the training scenarios to be relatable at times while not at others, accounting for the 25 responses among the 21 participants. Additionally, it is worth mentioning that all 3 faculty interview participants found the training scenarios to be unrelatable. Similar trends were observed in the user experience questionnaire. Participants were asked to evaluate the relatability of the stories as well as the believability of the stories’ endings at the 2nd, 4th, and 6th week marks of their study participation, resulting in a total of 209 responses. The majority of participants provided responses within moderate ranges for both questions, with 36% (76/209) finding the stories somewhat relatable and 32% (67/209) finding the endings somewhat believable (see Table 3). On average, faculty members tended to rate the scenarios as less relatable and the endings as less believable compared to the other participants.

**Repetitiveness**

Among the interview participants, 54% (12/22) found the training sessions to be repetitive. In certain aspects, repetition was intentionally incorporated into HTC as a way to give users repeated practice in thinking more flexibly and exposure to different scenarios. However, some found the repetitive writing style of the scenarios resulted in predictable outcomes, reducing the effectiveness of the training sessions.

*But at some point it just became very repetitive and instead of getting something out of it, I just saw a pattern and how to fix it.* [Undergrad 46]

Others experienced boredom due to the repetition of similar stories and themes.

*There were only so many scenarios or stories. And even though they were still very much applicable to my life, I could anticipate what those stories were gonna be and I just got bored.* [Grad 35]

**Length**

Of the interview participants, 82% (18/22) commented on the length of the training sessions.

Each microdose session consisted of 10 CBM-I scenarios, along with pre-and post- intervention
questionnaires. Researchers estimated that completing one microdose session would require roughly 5-10 minutes. Among the 18 participants that discussed session length, the majority (13/18, 72%) found the training sessions to be an appropriate length, although some expressed a preference for shorter sessions on occasion.

I would've liked for it to be shorter, but at the same time I can understand that the longer you're thinking about it, the more realistically and detailed you can imagine your situation. And with this topic, you really kind of should be spending a little bit more time on it than you want to. So I think it was an appropriate length, even though at times it was not fun. [Staff 41]

I thought they were a good length because I didn't have to worry about starting one and not being able to finish it. [Undergrad 35]

However, 28% (5/18) of participants that discussed length considered them too long, potentially impacting engagement.

The session was just a bit long and I think maybe that was one reason that kept me from doing it as often. I just dreaded doing it because maybe it was too long. [Grad 26]

For comparison/triangulation, a survey question regarding session length was administered to all participants during their 2nd week of study participation. Participants were asked how they felt about the session length and 92 responses were received with 50% (46/92) of participants indicating that the sessions were of the right length and 45% (41/92) of participants indicating that the sessions were too long (see Table 3).

Write-your-own

Of the interview participants, 64% (14/22) of interview participants provided feedback on the write-your-own scenario prompt. Some participants enjoyed the write-your-own scenario because it gave them a chance to apply their learning to real life situations.

I think that was my favorite part because it kind of forced me to think a little more deeply about the modules I was doing rather than just answering the questions. [Undergrad 35]
Conversely, others found the write-your-own scenario to be frustrating because it took too long or felt redundant. Additionally, concerns about privacy surfaced, particularly among faculty and staff members, regarding inputting personal stories into the app.

Crafting my own stories just irritated me...the homework aspect of it felt icky. But then also I didn't really want to because this is where I work. I didn't want to put scenarios in there...I think the exercise of thinking through my own scenarios would've been helpful if it wasn't writing them into the app. [Staff 39]

Theme 3: Barriers to Using the App

Lack of Understanding of Intervention Format or Rationale

Among the interview participants, 55% (12/22) expressed a lack of understanding of how to engage with the training sessions or the purpose of the intervention. Participants were given the flexibility to choose when and how many training sessions they completed on a given day, although it was recommended that they complete at least 2 microdose sessions. However, some participants found the structure of the app and study expectations to be ambiguous. Similarly, some participants were confused about the nature of the CBM-I training and did not understand the purpose of the tasks they were being asked to complete or how to engage with the stories.

And at least at the beginning of the study I was like, what am I required to do? What are the tasks for today? [Grad 52]

And so the very first survey, which was kind of longer, I thought, wait, is this like a reading comprehension survey? Like I really didn’t know what it was. And so I would read the scenario and then it would say essentially, what do you think would be happening next? And I'm like, I don't know. [Faculty 8]
After each training scenario, participants were asked questions about the story and were supposed to answer based only on information given in the scenario. The purpose of this was to practice thinking in new ways about different situations. However, some participants found it unclear whether they were expected to answer questions based on information given in the story or based on their current experiences. This lack of understanding led some to feel confused when their answers were marked as incorrect.

*I just do not understand why it would say wrong when both of them [answer choices] could be logical, and even sometimes I chose one that seemed like the thing that I would make happen or that was happening in my life, and it'd be wrong…why am I being told I'm wrong? [Faculty 1]*

**Forgetfulness**

Of the interview participants, 55% (12/22) cited forgetfulness as a reason for not using the app.

*I would usually just forget…I'm like already stressed out and then I'd scroll past it [the app], and I'm like, oh my god, I've forgotten to do that. [Undergrad 25]*

Although the app did send push notifications, some participants encountered issues receiving notifications, which contributed to forgetfulness.

*If I don't get the notification on my phone and if I'm out, I'm not going to remember this. [Grad 31]*

Some participants who mentioned not using the app due to forgetfulness expressed a preference for different types of notifications, such as calendar reminders rather than push notifications.

*It would've been better and more useful if it [reminders] would just populate [in my phone calendar] because I live or die by what's in my calendar. And if it's not on my calendar, I forget because there's just so much that's going on all day long. [Staff 39]*
Theme 4: Use Patterns

Used Routinely

Of the interview participants, 68% (15/22) used the app as part of their daily routine. These individuals found it most helpful to stick to a consistent schedule of using the app at the same time every day, rather than varying the time day by day or using the app specifically during moments of stress or anxiety.

I mostly did it at the same time each day just because of how my schedule works.

[Faculty 8]

I would do it in the morning, and then right before I went to bed. I don’t really remember doing it specifically if I was super anxious about something [Grad 41]

Used When Reminded

Similarly, 50% (11/22) of interview participants used the app because they received a notification or reminder.

I also liked that the app would send reminders, like you could set the time for every day that it would send reminders and as someone who also has been diagnosed with ADHD, like that’s the only way I can get work done is if I have reminders. [Grad 35]

I liked the feature where you could set a reminder to do them [training sessions] because that was like how I remembered to do it most days. [Undergrad 28]

Chosen Academics/Work/Career Development Domain

Participants also shared which stressor domain they used most often. Of the 8 domains to choose from, 82% (18/22) of interview participants stated they found the academics/work/career domain to be most important.

I feel like I chose the work one most of the time because that is honestly where most of my anxiety comes from. [Staff 45]
I tried to hit them all [domains], but the biggest source of stress in my life did happen to be graduate school...And so the graduate school one is definitely where I focused. [Grad 11]

Theme 5: Suggestions for Improvement

Feature Suggestions

Of the interview participants, 86% (19/22) suggested additional features to improve the apps functionality or enhance engagement. Among those participants, some suggested a progress tracking feature that would allow them to view their pending and completed tasks.

If you want people to take three surveys a day or something, you can show like a little Domino's pizza loading bar and be like, you've already done 33%, you have two more left. Or like a checkbox or something for people to tangibly be like, yeah, I did this. [Staff 41]

One thing that I just kind of thought of is I used the Duo Lingo app to help me to practice my French, and it tells me how well I've done recently. And if I have a streak of days in a row, it lets me know, hey, you've got this many days in a row. And it wasn't always clear to be able to see what my activity [in the HTC app] had been recently. [Faculty 4]

Others focused more on enhancing the customization and personalization of the app through tailored scenario content, adjustable training lengths, and flexible notifications. For example, some participants noted that certain scenario content was not relatable to their specific life stage.

I don't know if you can collect some sort of information on the individual who's taking it, but a lot of times, it [the scenarios] would be something like, you're going on a Tinder date, and I'm married with two kids. That's pretty hard for me to imagine. [Staff 41]

Consequently, these participants often suggested personalizing the content by filtering scenarios specific to their demographic to enhance relevance.
If there was some way to like screen participants and tailor the questions, I don’t know, but if it was tailored to something that was more relatable to each participant, I feel like they could get more out of this program. [Undergrad 46]

Additionally, some expressed a desire to customize the length of the training sessions. I hoped it was longer when I had a bad day. And I hoped it was shorter when I was having a busy day. Maybe it can be helpful if I can adjust the length. [Grad 13]

Lastly, some participants suggested enhancing the notification system and expressed wanting more customizable notification preferences, including frequency and delivery method. I wish that just the average number of notifications was a little lower. And maybe people can opt into more or opt out of more. [Faculty 4]

Maybe something that should be integrated is emails along with the notifications...if you’re on your computer and not on your phone, maybe the different notifications would help. [Grad 41]

**Domain Suggestions**

Finally, 73% (16/22) of interview participants suggested additional domains or ways to improve the existing ones. Of those 16 participants, 4 (25%) provided specific suggestions to improve the discrimination domain. The discrimination domain differed from the other 7 domains in that it did not provide specific scenarios for CBM-I training. Instead, it provided tools and resources for: 1) reporting discrimination or harassment, 2) self-care and wellness, 3) allyship (supporting others who have experienced discrimination, marginalization, or harassment), 4) community support, and 5) activism. Some participants appreciated how this domain was set up, while others mentioned wishing that this domain offered CBM-I training, similar to the other domains. I think I clicked on the discrimination topic one time…but it just brought me straight to links for like resources. So I think expanding that topic could be good because with more and more DEI conversations and stuff like that, people get really uncomfortable talking about it. So I think having something more in that area, like the resources were good and
I’m grateful that they were there, but I think having something similar where you could kind of work out your ideas there too. [Staff 41]

But if I click that category, then it shows that there’s like this, this, and this kind of resource, rather than showing up the scenarios like it did for the other categories, and I felt that a little unhelpful…for instance, I always struggle to think that my English is not a problem, and that’s also true, but like sometimes when I speak some wrong sentences, I become very small and start to be very conscious about that. But people around me are more likely to be friendly and do not really care about that. So I hope there was some scenarios that helped me think that like all of their reactions are not really a discrimination. I don’t, I know it can be very tricky, like it can be the app justifying there’s no discrimination, but still I want something that can help or support mentally. [Grad 13]

Others suggested additional resources and tools within the discrimination domain that they wished were provided. This included adding resources specific to political and religious discrimination.

The discrimination that I faced that made me go on was political discrimination…So I went on there to see if I could get any help with that. But the resources I don’t think were really applicable to me in that sense. [Grad 26]

I don’t know if you guys had something about religion…I think that’s an important aspect.

Maybe that could go with the discrimination. [Grad 49]

Additional content recommendations were proposed for other domains. These include incorporating more scenarios tailored to an older demographic, such as addressing anxiety related to retirement, making friends later in life, dealing with a child's college rejection, coping with the loss of a family member, or managing the estate of a deceased parent.

In addition to suggestions made to the existing domains, other participants provided suggestions to add new domains, such as a domain specific to international students.
I don't recollect right now but I don't know if there was a separate section that was directed towards international students...because there's so many international students and we have like our own set of problems, which sort of loosely relate to discrimination, but then also being homesick, being away from family, which weren't there. [Grad 35]

Discussion

This study used mixed methods to explore users’ experiences with the HTC platform and collect feedback on acceptability and ways to improve the app, which may also inform future designs of DMHIs more broadly. Qualitative data was collected via user feedback interviews and quantitative data via user experience questionnaires. Specifically, user experience questions regarding the effectiveness of the training program, the relatability of the training scenarios, and the length of the training sessions were analyzed to help validate and enrich the qualitative results (see Table 3). Results showed that the quantitative findings aligned with the qualitative data as the majority of both survey and interview participants found the program to be at least slightly or somewhat helpful in reducing or managing their anxiety or stress. Similarly, both survey and interview data yielded varied opinions on the relatability of the stories and found the length of the training sessions to either be an appropriate length or slightly too long. Overall, results supported the acceptability of the program.

Among those who did not perceive the HTC program as effective, there seemed to be a gap in the relatability of content and alignment with needs/goals, potentially influencing their perceived efficacy. For example, some participants expressed not experiencing significant stress or anxiety prior to using the app, which could explain their limited observations of changes in their stress levels. Given participants were not required to meet a specific anxiety threshold to join the study, it is possible that other participants had similar experiences, though the mean baseline OASIS score indicated significant anxiety on average (a score of 8 is typically considered a clinical cutoff where most users would meet criteria for an anxiety disorder...
diagnosis and our sample mean was ~6.5). Moreover, one participant reported having other mental health conditions outside of anxiety that the app is not targeted to treat. Some participants also reported a lack of comprehension regarding the intervention's purpose, and a few stated they downloaded the app expecting it to be more meditative in nature. These findings underscore the importance of providing a clear introduction to the program's format and rationale during enrollment, allowing the user to understand the overall structure of the app and determine if it is a good fit for them. However, it is important to note that HTC does provide written and video explanations regarding the program rationale, learning goals, and ways to use the app during the first training session. This raises questions about whether users are effectively absorbing the introductory content or if changes are needed to present the program overview more effectively. Further work is encouraged in this area.

While many users did report perceived benefits from the program, qualitative interviews were conducted with an emphasis on gathering feedback for future app versions. Therefore, this discussion primarily centers on suggestions to enhance future iterations of HTC and DMHIs more broadly. For example, offering explicit instructions on program expectations and guidelines for recommended use in multiple places may enhance user engagement, given this information seems to be readily missed or skimmed. Consistent with prior qualitative research, insufficient task explanations have been shown to deter users from continuing with digital health interventions [39]. Therefore, it is essential to present instructions clearly, and it may help to use various mediums [40]. HTC incorporates text-based and video instructions; to enhance clarity regarding task expectations, our users specifically suggest implementing a daily progress tracker, which could include gamification elements that might increase user motivation and engagement [41] while also clarifying what tasks need to be completed each day. Consistent with other studies, our findings also suggest increasing app personalization would be particularly beneficial in the context of CBM-I scenario content, as some users expressed dissatisfaction with the relatability of the stories [42, 43]. Given CBM-I relies on users imagining
themselves in various situations to promote more flexible thinking, creating more personally
relevant content may enhance users’ ability to engage with the exercises more effectively.
Therefore, we suggest future research should explore increasing personalization by using
demographic information to filter content based on factors such as age, relationship status, and
occupation, or assessing the types of interactions or places where each person tends to
experience anxiety, or the particular unhelpful interpretations they tend to make. We would also
like to highlight that despite our efforts to recruit faculty members, engaging them proved
challenging, resulting in a small sample size for this study. However, it is worth noting that all
three faculty participants indicated that the study content was not relatable. Therefore, we
believe increasing app personalization may be especially beneficial for this demographic.
Along with desiring personalized training content, users also value flexibility and autonomy,
particularly expressing a desire for more customizable notification and content delivery options.
Notifications are of particular importance given a main barrier reported to engaging with the
HTC platform was forgetfulness. Additionally, many participants reported using the program
specifically when reminded and appreciated that HTC allowed them to schedule push
notifications at specific times. To further enhance notification customization, we recommend
offering users a range of delivery options such as push notifications, emails, or calendar
reminders. Other research has shown that sending notifications containing tailored messages
may increase engagement [44].
Similarly, some participants expressed the desire to customize content type and length. For
example, the HTC program provided three CBM-I training scenario types: long scenarios, short
scenarios, and write-your-own scenarios. Some participants reported receiving a write-your-own
training scenario, but not having the time or bandwidth to complete it at the moment, leading
to close the app entirely. They wished for the option to skip or choose certain scenario
types as needed. Additionally, some participants expressed a desire to customize the length of
microdose sessions, stating that sometimes they wanted to complete more or fewer than 10
training scenarios in one session. Although allowing participants to customize the type and length of training sessions may increase app engagement and user satisfaction, there is a concern that increasing customizability of training scenario type and length may sacrifice program effectiveness. Further work is needed to determine the best balance between increasing training flexibility and maintaining intervention effectiveness (e.g., ensuring users complete an adequate dose of training).

Last, offering more diverse training content is likely important for increasing engagement and accessibility. Research indicates that repetitive training content can lead to boredom and disinterest [45, 43]. Incorporating a wider variety of scenario content and formats, including those suggested by users, such as scenarios relevant to international students, may improve acceptability and adherence. We recommend continuing to involve users in the co-design process for future iterations of HTC, particularly to develop more relevant scenarios for marginalized communities. This approach ensures that the content remains engaging and inclusive [46].

Overall, our findings suggest that HTC is viewed as somewhat acceptable and effective at reducing stress and anxiety, and we provide suggestions for increasing user engagement and app satisfaction across HTC and other mHealth platforms. To enhance engagement, we recommend implementing a robust notification system and providing clear instructions on study tasks and app use expectations through interactive modules and progress trackers. Additionally, we recommend further personalizing content and increasing customizable features to boost user satisfaction and adherence. Lastly, we emphasize the importance of involving users in the co-design process of DMHIs to ensure content remains relevant and meaningful to the intended audience.

**Limitations**

Several limitations should be considered in interpreting this interview and survey feedback. In some cases, individual qualitative interviews of participants took place a few months after their
initial use of the intervention. This gap may have biased participants' reflections regarding their experience with HTC or its effectiveness. To address this, future iterations of the intervention could schedule post-treatment interviews more promptly after the completion of the program to gather participants' more immediate impressions of the program, as well as some interviews later to determine longer term impacts of the program.

Furthermore, while the sample size is consistent with related studies in the literature [20, 30], it likely does not sufficiently represent the experience of all participants, especially considering the potential for sampling biases among the subset of users who agreed to take part in the post-intervention interviews. Moreover, we did not screen users for elevated anxiety before participation, so some participants did not meet a standard cutoff for clinical anxiety. This could skew participants' perception of the app's effectiveness if they do not struggle with anxiety or if they struggle with a different mental health condition that the app is not targeted to treat. In the future, screening for individuals who are high in trait, social, or generalized anxiety could provide a more appropriate sample to study the program's effectiveness (though we also value letting users choose whether they believe the program is something they want and need, as we did in the current study). We also did not require participants to stop other interventions while using HTC; this was done to increase external validity but we do not know how the experience of using HTC varies based on whether it is used on its own or as an adjunct to other forms of care. Finally, while the coders had varying levels of familiarity with HTC and considered diverse perspectives in the interviews, there is a possibility of reflexivity within the coding process.

**Conclusions**

We collected qualitative interview data and quantitative survey data to better understand user experiences of the HTC platform and suggestions for its improvement. Themes that appeared in over half of the interviews regarding the effectiveness of HTC included an ability to apply new coping strategies in daily life and a perceived change in anxiety levels post-treatment. The most common response indicated that HTC was somewhat helpful at reducing anxiety. There were
also many suggestions for improvements to enhance relatability and personalization of training
content to better sustain engagement, along with a need to improve understanding of the
program’s format and rationale, and easy ways for users to be reminded to use the program.
Together, these findings point to the importance of iterative intervention development and
regularly partnering with users to co-design DMHIs so they are acceptable, feasible, and can
meet the needs and preferences of diverse users.

References

1. Meeks K, Peak AS, Dreihaus A. Depression, anxiety, and stress among students,
2. Kazdin AE. Addressing the treatment gap: A key challenge for extending evidence-
based psychosocial interventions. Behaviour Research and Therapy. 2017
Jan;88(88):7–18.
3. Weissinger G, Ho C, Ruan-Iu L, Van Fossen C, Diamond G. Barriers to mental health
services among college students screened in student health: A latent class analysis.
4. Wagner B, Snoubar Y, Mahdi YS. Access and efficacy of university mental health
Mental Health and Well-Being among University Faculty: A Narrative Review. J Clin
Med. 2023 Jun 30;12(13):4425. doi: 10.3390/jcm12134425. PMID: 37445459; PMCID:
PMC10342374.
6. Wright JH, Mishkind M. Computer-Assisted CBT and Mobile Apps for Depression:
Mental Health Interventions for Depression, Anxiety, and Enhancement of Psychological


24. Livermon S, Michael A, Petz KD, Toner ER, Teachman B. A Mobile-Based Intervention to Reduce Stress, Anxiety, and Depression Among University Students, Faculty, and Staff: Qualitative Study on Users’ Experiences [Internet]. OSF; 2023. Available from: osf.io/a46hb


34. Delve Tool [Internet]. [cited 2024 Jul 17]. Available from: https://delvetool.com/


37. Livermon S. A Mobile Intervention to Reduce Anxiety Among University Students, Faculty, and Staff: Mixed Methods Study on Users’ Experiences [Internet]. OSF; 2024. Available from: osf.io/w8jzg

38. Fusch PI, Ness LR. Are We There Yet? Data Saturation in Qualitative Research [Internet]. NSUWorks. 2015. Available from: https://nsuworks.nova.edu/tqr/vol20/iss9/3/

Modes of Delivery, Facilitators, and Barriers. Journal of Medical Internet Research. 2020


40. Berry N, Lobban F, Bucci S. A qualitative exploration of service user views about using
digital health interventions for self-management in severe mental health problems. BMC
PMCID: PMC6341527.

41. Bakker D, Kazantzis N, Rickwood D, Rickard N. Mental Health Smartphone Apps:
Review and Evidence-Based Recommendations for Future Developments. JMIR Mental
Health [Internet]. 2016 Mar 1;3(1):e7. Available from:
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4795320/

acceptability of a digital mental health platform incorporating a virtual coach: The good,
the bad, and the opportunities. Health Informatics Journal. 2021
Jan;27(1):146045822199487.

43. Balaskas A, Schueller SM, Cox AL, Doherty G. Understanding users’ perspectives on

44. Bidargaddi N, Almirall D, Murphy S, Nahum-Shani I, Kovalcik M, Pituch T, et al. To
Prompt or Not to Prompt? A Microrandomized Trial of Time-Varying Push Notifications to
Increase Proximal Engagement With a Mobile Health App. JMIR mHealth and uHealth.
2018 Nov 29;6(11):e10123.

Centered App Adaptation of a Low-Intensity E-Mental Health Intervention for Syrian

46. Porche MV, Folk JB, Tolou-Shams M, Fortuna LR. Researchers’ Perspectives on Digital
Mental Health Intervention Co-Design With Marginalized Community Stakeholder Youth
Supporting Information

S1 Figure. Thematic Maps (initial and revised to only include the most saturated codes).

S1 Text. Standards for Reporting Qualitative Research (SRQR) Checklist.

S2 Text. Semi-structured Interview Guide.