Original Paper
Russell Fulmer, PhD; Angela Joerin, MS, LLP; Breanna Gentile, PhD; Lysanne Lakerink; Michiel Rauws, MBA

Using Integrative Psychological Artificial Intelligence to Relieve Symptoms of Depression and Anxiety in Students

Abstract
Background: Approximately 20 million college students suffer from mental illness in the United States alone [1]. More than 50% of college students report experiencing symptoms of depression and anxiety that impact daily functioning within the last year [2]. Despite a clear need for clinical services, up to 75% of college students do not access adequate mental health care [3]. With a growing desire for on-demand services that engage students and reduce stigma, web and mobile based mental health interventions offer a scalable solution. Mental health care solutions such as computer-assisted therapy (CAT) have been shown to be a less-intensive and more cost-effective method to deliver empirically-validated treatments for depression and anxiety [4]. Although traditional in-person treatment remains the standard of care for those with clinical levels of depression, preliminary studies suggest that self-help computer-based cognitive and behavioral interventions produce similar outcomes [5], and are efficacious in the treatment of sub-threshold mood disorders. One study revealed that nearly 70% of patients expressed interest in using mobile health apps to self-monitor and self-manage their mental health [6]. Early evidence suggests that patients open up more while using a mobile health app than during face-to-face therapy. With one app for patients suffering from suicidal thoughts, more subjects reported suicidal ideation using the app than they did on the traditionally administered PHQ-9 [7]. Psychological AI agents delivering cognitive-behavioral therapy (CBT) have been shown to be a feasible, engaging, and effective solution for reducing symptoms of depression and anxiety in college students [8]. However, the efficacy of using psychological AI agents to deliver integrative mental health care, including CBT, requires further exploration. Although CBT is one of the most effective methods for treating anxiety and depression, evidence shows that alternative forms of therapy lead to equally successful outcomes [9]. Applying an integrative approach to therapy for treating patients with depression has been linked to greater posttreatment improvement and more clinically significant change than those receiving cognitive therapy [10].

Objective: The objective of this study was to assess the efficacy of using an integrative psychological Artificial Intelligence Chatbot (AI), Tess, to reduce symptoms of depression and anxiety in an engaging way. Tess was tailored by the research team to best meet student needs and designed to deliver personalized conversations based on the emotions and mental health concerns expressed by participants. Emotions are expressed and identified in a variety of ways. Tess focuses on language as the most explicit form of communication with the proposition that communication between people reveals individual conceptualizations of specific emotions [11]. As an integrative
psychological AI, Tess was designed to aggregate and analyze the emotions and concerns expressed to build up an understanding of each participant's emotional profile, and in turn respond with the most appropriate intervention, including crisis resources.

Methods: Participants were recruited using a flyer posted through social media outlets targeting students from 15 universities across the United States. Inclusion criteria included current enrollment at a university in the United States, age 18 and over (screened at the first level via checkbox confirmation) and able to read English (implied). To guard against compromise, for example from malicious bots, all potential participants were sent an email requesting that they respond using their university email denoting their confirmation. Confirmed participants were randomized via computer algorithm that automatically generated a number between one and three. Participants with number one were sent a link to the National Institute of Mental Health's ebook on depression among college students, after completion of online baseline questionnaires. Participants with numbers two and three were allocated to receive a direct link to begin chatting with Tess via an instant messenger app. Because the randomization allocation occurred algorithmically, allocation concealment was in place. However, the condition to which each participant was allocated was not masked for the service providers (Tess). After approximately two weeks (Group 2) or four weeks (Group 1+3), participants were contacted again to complete a second set of questionnaires online. Participants were offered a prorated incentive of US $20 for completion of both assessments. Since this trial involved a nonclinical population of college students, it was considered exempt from registration in a public trials registry.

Results: A statistically significant difference was found on the Positive and Negative Affect Schedule (PANAS) between the control group and group 1, which had access to Tess each day for two weeks. The P value of 0.028 is statistically significant at an alpha level of 0.05 and confirms that Tess did impact scores. A statistically significant difference was found on the PHQ-9 between the control group and group 1, which had access to Tess each day for two weeks. The P value of 0.027 is statistically significant at an alpha level of 0.05. This finding confirms that Tess was helpful in decreasing depressive symptoms. A statistically significant difference was found on the GAD-7 between the control group and both groups 1 and 2. For group 1, the P value of 0.045 is statistically significant at an alpha level of 0.05. For group 2, the P value of 0.021 is statistically significant at an alpha level of 0.05. This finding supports the hypothesis that Tess would be helpful in decreasing anxiety symptoms.

Conclusions: Principal Results
It is notable that our control group was selected at random and by chance had higher levels of depression and anxiety. An unintended consequence of our study is that the control group had higher levels of depression and anxiety at four weeks than at baseline. One possible explanation for this outcome is that the e-book may have made the control group more aware of symptoms of depression and anxiety without offering interventions to reduce symptoms. This further demonstrates the value in receiving support through Tess. All participants were offered mental health support through our Psychological AI agent upon completing the study.
Comparisons With Prior Work
Aligned with results from a previous study, using Tess was associated with a significant reduction in depression and anxiety as measured by the PHQ-9 and GAD-7, respectively. The effect size (Cohen $d=0.68$) for depression was moderate and greater than previously published studies [8, 26-28] that measured the efficacy of using alternative mobile app interventions to relieve symptoms of depression. The current study included two test groups to evaluate differences in symptom reduction based on two to four week intervention periods. The effect size for reduction of symptoms is aligned with that found by Fitzpatrick et al when delivering a CBT based conversational agent to college students in the United States. The greater effect size found in the current study may be due to the integrative mental health approach applied by Tess to deliver more personalized interventions. With a growing demand for scalable solutions that deliver more cost-effective mental health support, it has been shown that computer-assisted therapy (CAT) is capable of delivering empirically-validated treatments for depression and anxiety [15]. Preliminary studies suggest that self-help computer-based cognitive and behavioral interventions produce similar outcomes to in-person treatment [16]. Clinical treatment outcomes have been higher for patients prescribed to use psychotherapeutic computer programs compared to programs that are delivered in a self-help format with no clinician involvement [30]. Previous studies suggest that individuals are more willing to disclose personal information to a psychological AI than to a “virtual therapist” purportedly operated by a human [7, 8, 31]. This was supported by feedback given by student participants who engaged on a more personal level with Tess for the current study. Students reported “I do somewhat feel like I'm talking to a real person,” “you're better than my therapist [who] doesn't necessarily provide specific ways I can better myself,” and that Tess was able to “coach [the participant] through a difficulty.” These comments reinforce the potential for psychological AI to remove barriers and stigma associated with traditional therapeutic methods. Trained by experienced mental health professionals, psychological AI is able to deliver interventions that mimic the therapeutic process including empathic reflection.

Trial Registration: N/A

Keywords: Artificial intelligence, integrative psychology, student depression, student anxiety, chatbot

Introduction
Approximately 20 million college students suffer from mental illness in the United States alone [1]. More than 50% of college students report experiencing symptoms of depression and anxiety that impact daily functioning within the last year [2]. Despite a clear need for clinical services, up to 75% of college students do not access adequate mental health care [3]. With a growing desire for on-demand services that engage students and reduce stigma, web and mobile based mental health interventions offer a scalable solution.
Mental health care solutions such as computer-assisted therapy (CAT) have been shown to be a less-intensive and more cost-effective method to deliver empirically-validated treatments for depression and anxiety [4]. Although traditional in-person treatment remains the standard of care for those with clinical levels of depression, preliminary studies suggest that self-help computer-based cognitive and behavioral interventions produce similar outcomes [5], and are efficacious in the treatment of sub-threshold mood disorders.

One study revealed that nearly 70% of patients expressed interest in using mobile health apps to self-monitor and self-manage their mental health [6]. Early evidence suggests that patients open up more while using a mobile health app than during face-to-face therapy. With one app for patients suffering from suicidal thoughts, more subjects reported suicidal ideation using the app than they did on the traditionally administered PHQ-9 [7]. Psychological artificial intelligence (AI) agents delivering cognitive-behavioral therapy (CBT) have been shown to be a feasible, engaging, and effective solution for reducing symptoms of depression and anxiety in college students [8]. However, the efficacy of using psychological AI agents to deliver integrative mental health care, including CBT, requires further exploration. Although CBT is one of the most effective methods for treating anxiety and depression, evidence shows that alternative forms of therapy lead to equally successful outcomes [9]. Applying an integrative approach to therapy for treating patients with depression has been linked to greater posttreatment improvement and more clinically significant change than those receiving cognitive therapy [10].

Thus, the objective of this study was to assess the efficacy of using integrative psychological AI, Tess, to reduce symptoms of depression and anxiety in an engaging way. Tess was tailored to best meet student needs and designed to deliver personalized conversations based on the emotions and mental health concerns expressed by participants. Emotions are expressed and identified in a variety of ways. Tess focuses on language as the most explicit form of communication with the proposition that communication between people reveals individual conceptualizations of specific emotions [11]. As an integrative psychological AI, Tess was designed to aggregate and analyze the emotions and concerns expressed to build up an understanding of each participant’s emotional profile, and in turn respond with the most appropriate intervention, including crisis resources.

This study compared outcomes from two to four weeks of using integrative psychological artificial intelligence (Tess), or an information control group (National Institute of Mental Health’s [NIMH] ebook) in a nonclinical college population. It was hypothesized that engaging in conversations with integrative psychological AI would lead to greater improvement in symptoms relative to the information control group.

Tess
Tess is a psychological artificial intelligence (AI) agent that delivers integrative mental health support, psycho-education, and reminders. Tess is available 24/7 through existing
Tess delivers mental health interventions that have repeatedly been shown to reduce symptoms of depression and anxiety. The primary form of therapy is cognitive behavioral therapy (CBT) [12], primarily due its strong evidence-base [13]. In addition to CBT, Tess also delivers a variety of similar, clinically proven therapies, dependent on both the emotions felt by the patient and the nature of the detected psychological issue. These include conversations based on the Transtheoretical Model [14], Emotionally Focused Therapy [15], [16], Solution-Focused Brief Therapy [17], Motivational Interviewing [18], Dialectical Behavior Therapy [19] and more. By interacting with Tess, the participant experiences the positive side-effects of journaling, which has been shown to increase the positive perception of experiences [20] and significantly improve self-efficacy [21].

Tess is capable of interpreting free text messages, alternatively users can opt for preselected responses that is the only option of many existing conversational agents. This enhances Tess’s capacity to deliver more personalized and integrative interventions. Just as therapists adjust their style to accommodate a client’s therapeutic preference over time, Tess works to build rapport and deliver interventions that best meet a user’s needs. The more messages exchanged between a user and Tess, the more Tess is able to identify therapeutic preferences and will be more likely to deliver interventions that the user will find helpful. For example, if a user responds positively to cognitive behavioral therapy (CBT) and negatively to self-compassion therapy, Tess will deliver more interventions rooted in CBT. During the study, Tess delivered interventions based in CBT emotionally focused therapy acceptance and commitment therapy, dialectical behavior therapy, self-compassion therapy, journaling, and interpersonal psychotherapy.

Tess replied in an empathic way appropriate to the participants’ inputted mood. For example, in response to endorsed loneliness, Tess replied “I’m so sorry you’re feeling lonely. I guess we all feel a little lonely sometimes” or it showed excitement, “Yay, always good to hear that!”. Specific content is sent to individuals depending on mood state. For example, a participant indicating that they feel anxious is offered in-vivo assistance with the anxious event. It is of note that Tess was disguised as “Zara” for this study to prevent bias in the unlikely chance that participants had been exposed to Tess through another initiative.

Methods

Recruitment

Participants were recruited using a flyer posted through social media outlets targeting students across 15 universities across the United States. Inclusion criteria included current enrollment at a university in the United States, age 18 and over (screened at the first level via checkbox confirmation) and able to read English (implied). To guard against compromise, for example from malicious bots, all potential participants were
CONFIDENTIAL

sent an email requesting that they respond using their university email denoting their confirmation.

Confirmed participants were randomized via computer algorithm that automatically generated a number between one and three. Participants with number one were sent a link to NIMH's ebook [22] on depression among college students, after completion of online baseline questionnaires. Participants with numbers two and three were allocated to receive a direct link to begin chatting with Tess via an instant messenger app. Because the randomization allocation occurred algorithmically, allocation concealment was in place. However, the condition to which each participant was allocated was not masked for the service providers (Tess). After approximately two weeks (Group 2) or four weeks (Group 1+3), participants were contacted again to complete a second set of questionnaires online. Participants were offered a prorated incentive of US $20 for completion of both assessments. Since this trial involved a nonclinical population of college students, it was considered exempt from registration in a public trials registry.

Measures
The Patient Health Questionnaire-9
The Patient Health Questionnaire (PHQ-9) [23] is a nine item, self-report questionnaire that assesses the frequency and severity of depressive symptomatology within the previous two weeks. It is one of the most widely used, reliable, and validated measures of depressive symptoms. Each of the nine items is based on the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) criteria for major depressive disorder and can be scored on a zero (not at all) to three (nearly every day) scale.

Generalized Anxiety Disorder-7
The Generalized Anxiety Disorder seven item scale (GAD-7) [24] is a valid, brief self-report tool to assess the frequency and severity of anxious thoughts and behaviors over the past two weeks. Based on the DSM-IV diagnostic criteria for GAD-7, the scores of all seven items range from zero (not at all) to three (nearly every day).

Positive and Negative Affect Schedule
The Positive and Negative Affect Schedule (PANAS) [25] is a 20-item self-report measure of current positive and negative affect. Half the items represent positive affect (i.e., interested, excited, determined), whereas half of the items are indicative of negative affect (i.e., hostile, scared, ashamed). Items are scored on a one (very slightly or not at all) to a five (extremely) scale, with higher scores representing higher affect. Positive and negative affect are summed independent of each other.

Results
Participant Demographics
Table 1 shows the demographic information and baseline scores on clinical variables for those with data from the entire sample (N=74). Participants were an average of 22.91 years old and over two-thirds female. The majority of participants were Asian (51.9%, 37 participants) and Caucasian (41.6%, 32 participants).
In the control group 67% (16) were female, 29% (7) were male, and 4% (1) participant identified as non-conforming. The average age for the control group was 22.5 years. The majority of the control group was Caucasian (46%, 11 participants). The remainder of the control group was Asian (33%, 8 participants), Other (13%, 3 participants), African American (8%, 2 participants).

Group 1 consisted of 17 females and 7 males (71% and 29%, respectively). The average age for Group 1 was 24.1 years. This group was mostly Caucasian (54%, 13 participants) and eleven participants identified as Asian (46%).

Group 2 consisted of 19 females and 7 males (73% and 27%, respectively). The average age for Group 2 was 22.19 years. This group has mostly Asian participants (69%, 18 participants) with 31% (8 participants) being Caucasian.

Participant Clinical Variables

Table 1 shows the scores for the two scales and for the subscales of the PANAS. In the control group the average PHQ-9 score was 8.17, for Group 1 it was 6.67, and for Group 2 it was 7.04. The GAD-7 had an average score of 9.46 for the control group, 6.71 for Group 1, and 7.5 for Group 2. The average positive affect scale of the PANAS for the control group was 22.13, for Group 1 it was 19.88, and for Group 2 it was 21.31. The average negative affect scale of the PANAS for the control group was 15.75, for Group 1 it was 13.08, and for Group 2 it was 14.38.

Figure 1. Participant recruitment flow.
Table 1. Demographic and clinical variables of participants at baseline.

<table>
<thead>
<tr>
<th>Scale, mean (SD)</th>
<th>Information control</th>
<th>Tess Group 1</th>
<th>Tess Group 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Depression (PHQ-9)</td>
<td>8.17</td>
<td>6.67</td>
<td>7.04</td>
</tr>
<tr>
<td>Anxiety (GAD-7)</td>
<td>9.46</td>
<td>6.71</td>
<td>7.5</td>
</tr>
<tr>
<td>Positive affect</td>
<td>22.13</td>
<td>19.88</td>
<td>21.31</td>
</tr>
<tr>
<td>Negative affect</td>
<td>15.75</td>
<td>13.08</td>
<td>14.38</td>
</tr>
<tr>
<td>Age, mean (SD)</td>
<td>22.5</td>
<td>24.1</td>
<td>22.19</td>
</tr>
<tr>
<td>Gender, N (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>16 (67%)</td>
<td>17 (71%)</td>
<td>19 (73%)</td>
</tr>
<tr>
<td>Male</td>
<td>7 (29%)</td>
<td>7 (29%)</td>
<td>7 (27%)</td>
</tr>
<tr>
<td>Non-conforming</td>
<td>1 (4%)</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
### Ethnicity, N (%)

<table>
<thead>
<tr>
<th>Ethnicity</th>
<th>Group 1 (N)</th>
<th>Control Group (N)</th>
<th>Information Control Group (N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>African American</td>
<td>2 (8%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Asian</td>
<td>8 (33%)</td>
<td>11 (46%)</td>
<td>18 (69%)</td>
</tr>
<tr>
<td>Caucasian</td>
<td>11 (46%)</td>
<td>13 (54%)</td>
<td>8 (31%)</td>
</tr>
<tr>
<td>Other</td>
<td>3 (13%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
</tr>
</tbody>
</table>

**Analysis**

A statistically significant difference was found on the PANAS between the control group and group 1, which had access to Tess each day for two weeks. The P value of 0.028 is statistically significant at an alpha level of 0.05 and confirms Tess did impact scores.

A statistically significant difference was found on the PHQ-9 between the control group and group 1, which had access to Tess each day for two weeks. The P value of 0.027 is statistically significant at an alpha level of 0.05. Figure 2 shows a multivariate analysis of covariance revealed a significant group difference on depression such that those in the test group significantly reduced their symptoms of depression over the study period as measured by the PHQ-9, while those in the information control group increased their symptoms of depression. The increase found in the control group is most likely due to inability to adjust for potential confounding variables (e.g., participants in the control group had a larger course load than participants in other groups). These findings confirm that Tess was helpful in decreasing depressive symptoms.

**Figure 2.** Change in depression by group (PHQ-9 score).

A statistically significant difference was found on the GAD-7 between the control group and both groups 1 and 2. For group 1, the P value of 0.045 is statistically significant at an alpha level of 0.05. For group 2, the P value of 0.021 is statistically significant at an
alpha level of 0.05. Figure 3 shows a multivariate analysis of covariance revealed a significant group difference on anxiety such that those in the test group significantly reduced their symptoms of anxiety over the study period as measured by the GAD-7, while those in the information control group experienced increased symptoms of anxiety. This increase can be explained similarly to the aforementioned depressive scores - inability to adjust for potential confounding variables. These findings support the hypothesis that Tess would be helpful in decreasing anxiety symptoms.

**Figure 3.** Change in anxiety by group (GAD-7 score).

Qualitative Results

Table 2 shows the difference between the information control group and the test group. At the end of the study all participants had to answer a few scale questions such as: “How satisfied were you overall?” and “How satisfied were you with the content?”. This table shows a significant difference between both groups. For example, 86% (n=43) is overall satisfied with Tess and only 60.8% (n=14) with the e-book. 80% (n=40) learned something new from Tess and 43.5% (n=10) learned something new from the e-book.

**Table 2.** Qualitative results to post survey questions.

<table>
<thead>
<tr>
<th>Post survey questions</th>
<th>% Control Group (N=24)</th>
<th>% Tess Group (N=50)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall satisfaction</td>
<td>60.8% is overall satisfied (n=14)</td>
<td>86% is overall satisfied (n=43)</td>
</tr>
<tr>
<td>Content satisfaction</td>
<td>65.2% is satisfied with content (n=15)</td>
<td>80% is satisfied with content (n=40)</td>
</tr>
<tr>
<td>----------------------</td>
<td>---------------------------------------</td>
<td>--------------------------------------</td>
</tr>
<tr>
<td>Extend emotional awareness</td>
<td>73.9% thinks the E-book facilitated emotional awareness (n=17)</td>
<td>86% thinks Tess facilitated emotional awareness (n=43)</td>
</tr>
<tr>
<td>Learned something new</td>
<td>43.5% learned something new (n=10)</td>
<td>80% learned something new (n=40)</td>
</tr>
<tr>
<td>Relevant to everyday life</td>
<td>65.2% thinks the information is relevant to their everyday life (n=11)</td>
<td>80% thinks the information is relevant to their everyday life (n=40)</td>
</tr>
<tr>
<td>More comfortable with therapeutic process</td>
<td>47.8% is more comfortable with the therapeutic process (n=11)</td>
<td>64% is more comfortable with the therapeutic process (n=32)</td>
</tr>
</tbody>
</table>

Figure 4 shows a thematic map of participants’ responses to the question “What was the best thing about your experience using Tess?” Two major themes emerged in respect to this question: process and content. In the process theme, the subthemes that emerged were accountability from accessibility (noted by 15 participants); the empathy that the bot showed (n=6); and the learning that the bot facilitated (n=11), which in turn was divided into further sub themes of emotions (n=3), general insight (n=8).

**Figure 2.** Thematic flow of participants’ most favored features while interacting with Tess.

![Thematic Flow Diagram]

Figure 5 illustrates a thematic map of participants’ responses to the question: “What was the worst thing about your experience with Tess?” Two themes emerged: process (n=31) and problems with content (n=19). The most common sub theme to emerge among the process violations related to the limitations in natural conversation (n=12)
and the bot not being able to understand some responses or getting confused when unexpected answers were provided by participants (n=11). Problems with content were described by 10 individuals, most of which related to not enough interactivity (n=7).

**Figure 2.** Thematic flow of participants' least favored features while interacting with Tess.

A total of 48 comments were received as feedback from participants, see figure 5. These comments were mostly positive and many expressed gratitude for the experience: “Based on our interactions I do somewhat feel like I'm talking to a real person and I do enjoy the tips you've given. In that sense, you're better than my therapist in that she doesn't necessarily provide specific ways I can better myself and problems.” Also, “I've been learning new things and I have some ideas on ways I can make small changes that could help me!”.
Discussion

Principal Results

Limitations
While this study included participants across 15 universities across the United States, the generalizability of results is limited. The methodology called for two test groups and one control group, making the number of participants per group more limited. Additionally, this study did not collect follow-up data to assess if benefits were sustained over time. Alternative to previous studies, the control group experienced a slight increase in symptoms of anxiety and depression, suggesting that the e-book was not a sufficient form of mental health support. Given that this group also reported higher levels of anxiety and depression upon starting the study, it's possible that exposure to the materials increased awareness of symptoms without providing ongoing treatment, leading to an increase over time. One study revealed that consumers of self-help books are more sensitive to stress and show higher depressive symptomatology [32]. Future studies should include control conditions that allow for a more direct comparison between delivery of services such as direct therapy, teletherapy, or interactive online courses. Error rate as expressed during qualitative feedback, is explained by the fact that this version of Tess was customized by the research team using the X2AI customization platform on a very limited budget.

Traditional therapeutic methods allow for emotional assessment on many different levels including facial expressions, body cues, tone of voice, and language. The psychological AI used in this study delivered interventions via conversation and therefore emotion identification was limited to language. It is unclear how much this limited the psychological AI’s assessment of emotion as language is the most readily available non phenomenal access people have to emotions. Assessing emotion through facial expressions [29, 30], appears unreliable due to the overlap of expressive characteristics among seemingly basic emotions, resulting in the taxonomy of facial expressions not adequately describing the taxonomy of emotions [11].

Comparison with Prior Work
Aligned with results from a previous study, using Tess was associated with a significant reduction in depression and anxiety as measured by the PHQ-9 and GAD-7, respectively. The effect size (Cohen $d=0.68$) for depression was moderate and greater than previously published studies [8, 26-28] that measured the efficacy of using alternative mobile app interventions to relieve symptoms of depression. The current study included two test groups to evaluate differences in symptom reduction based on two to four week intervention periods. The effect size for reduction of symptoms is aligned with that found by Fitzpatrick et al when delivering a CBT based conversational agent to college students in the United States. The greater effect size found in the current study may be due to the integrative mental health approach applied by Tess to deliver more personalized interventions. Additionally, the content used to create Tess conversations was derived from written transcripts that allowed participants to respond
With free text, versus predominantly using buttons to receive videos and other resources on a timely basis as the chatbot of Fitzpatrick et al did.

With a growing demand for scalable solutions that deliver more cost-effective mental health support, it has been shown that computer-assisted therapy (CAT) is capable of delivering empirically-validated treatments for depression and anxiety [15]. Preliminary studies suggest that self-help computer-based cognitive and behavioral interventions produce similar outcomes to in-person treatment [16]. Clinical treatment outcomes have been higher for patients prescribed to use psychotherapeutic computer programs compared to programs that are delivered in a self-help format with no clinician involvement [29].

Previous studies suggest that individuals are more willing to disclose personal information to a psychological AI than to a “virtual therapist” purportedly operated by a human [7, 8, 31]. This was supported by feedback given by student participants who engaged on a more personal level with Tess for the current study. Students reported “I do somewhat feel like I'm talking to a real person,” “you’re better than my therapist [who] doesn't necessarily provide specific ways I can better myself,” and that Tess was able to “coach [the participant] through a difficulty.” These comments reinforce the potential for psychological AI to remove barriers and stigma associated with traditional therapeutic methods. Trained by experienced mental health professionals, psychological AI is able to deliver interventions that mimic the therapeutic process including empathic reflection.

Conclusions
This study reveals that psychological AI appears to be a feasible, engaging, and effective way to deliver integrative mental health support for some 10 million college students in the United States. The results support and expand upon findings from a previous randomized control trial [8] and demonstrates that psychological artificial intelligence has the potential to reduce symptoms of depression and anxiety by delivering CBT based interventions in the form of conversations.

Acknowledgements
Russell Fulmer, PhD for academic advisory contributions; Angela Joerin, MS, LLP for project management; Bre Gentile, PhD for statistical analysis; Lysanne Lakerink for research design and implementation; Michiel Rauws, MBA for review and approval of the manuscript; X2AI for providing a budget for the student reimbursements, and giving the research team free access to their customization platform and chatbot.

Conflicts of Interest
None declared.

Abbreviations
JMIR: Journal of Medical Internet Research
AI: Artificial Intelligence
PHQ: Patient Health Questionnaire
GAD: Generalized Anxiety Disorder
PANAS: Positive and Negative Affect Scale
CBT: Cognitive Behavioral Therapy
CAT: Computer-Assisted Therapy
NIMH: National Institute of Mental Health

References


8. Fitzpatrick KK, Darcy A, Vierhile, M. Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): A Randomized Controlled Trial. JMIR Mental Health; 2017;4(2);e19. [doi: 10.2196/mental.7785]


23. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613. [Medline: 11556941]


