Improving Smartphone Users’ Affect and Wellbeing with Personalized Positive Psychology Interventions

Sooyeon Jeong
MIT Media Lab
Cambridge, USA
sooyeon6@mit.edu

Cynthia Breazeal
MIT Media Lab
Cambridge, USA
cynthiab@media.mit.edu

ABSTRACT
We developed a smartphone application that detects users’ affect and provides personalized positive psychology interventions in order to enhance users’ psychological wellbeing. Users’ emotional states were measured by analyzing facial expressions and the sentiment of SMS messages. A virtual character in the application prompted users to verbally journal about their day by providing three positive psychology interventions. The system used a Markov Decision Process (MDP) model and a State-Action-Reward-State-Action (SARSA) algorithm to learn users’ preferences about the positive psychology interventions. Nine participants were recruited for an experimental study to test the application. They used it daily for three weeks. The interactive journaling activity increased participants’ arousal and valence levels immediately following each interaction, and we saw a trend toward improved self-acceptance levels over the three week period. The interaction duration increased significantly throughout the study as well. The qualitative analysis on journal entries showed that the application users explored and reflected on various aspects of themselves by looking at daily events, and found novel appreciation for and meanings in their daily routine.

Author Keywords
Positive psychology; mobile technology; psychological wellbeing; expressive writing therapy;

ACM Classification Keywords
H.1.2 User/Machine Systems: Software psychology

INTRODUCTION
Major depressive disorder is the leading cause of disability in the U.S. for ages 15-44 [2] and affects 14.8 million American adults, or about 6.7 percent of the U.S. population age 18 and older [15]. Various treatment methods, such cognitive behavioral therapy (CBT), interpersonal therapy, and medications have been developed in order to treat the disorder. However, individual psychotherapy is still not yet widely accessible to the majority of people who need psychological interventions [14,31].

In order to increase the accessibility of psychotherapy, many smartphone or computer-based applications and wearable sensors have been developed [7,8,16,17,19,21,23,32]. Most of these applications target patients who already have been diagnosed with a mental disorder and focus on monitoring their mental and/or physiological states, or providing guidance for CBT.

In this paper, we present a smartphone application that combines expressive writing therapy and positive psychology interventions in order to enhance smartphone users’ psychological wellbeing and affect. Expressive writing therapy has been shown to reduce depression and anxiety, increase physiological arousal, and decrease long-term health problems [3,11,18,22]. Positive psychology interventions can reduce depressive symptom levels for those diagnosed with major depressive disorder, and increase subjective and psychological wellbeing for patients with depression as well as for people who are not diagnosed with any mental disorder [1,8,21,22,27]. With the application, smartphone users verbally responded to a virtual character’s positive psychology intervention prompts. The virtual agent was not presented as a conversational partner but rather as a facilitator and a helper for users to reflect with, and interactively journal about their day.

The system analyzed users’ facial expressions and the sentiment of their SMS messages in order to learn which intervention prompts increased users’ affect and engagement over multiple interactions. We compared smartphone users’ affect immediately before and after using the application. Their psychological wellbeing levels were also measured over three weeks to evaluate any long-term change. Markov Decision Process (MDP) [24] and State-Action-Reward-State-Action (SARSA) [25] algorithms were used to learn the user’s preference for the interventions.

We hypothesized that (1) smartphone users’ affect and psychological wellbeing would increase after using the interactive journaling application daily for three weeks, and (2) user engagement with the application would increase...
over time as the application learned to select the most appropriate interventions for the users. The efficacy of the interactive journaling application in increasing users’ psychological wellbeing was evaluated by running a three-week long experimental study. The verbal journal entries were audio recorded on the device and then retrieved by the experimenter at the end of the study for qualitative analysis.

METHODS

Participants and Procedure

Nine participants (three male and six female, age M=28.3, SD=6.58) who owned Android devices with API 4.1 or higher were recruited from MIT campus via email advertisements. Once consented, the participants were asked to fill out the pre-survey and had the interactive journaling application installed on their mobile device for the study. Participants were asked to use the interactive journaling application every day for three weeks. The app notified the participants to use the application at 9PM every day, but did not provide constraints on the time or the length of app usage.

Participants were asked to fill out questionnaires on their mood [20], perceived stress [6], affect balance [5] and psychological wellbeing [26] for pre-, mid- and post-tests (Table 1). The pre-questionnaire was administered at the beginning of the study. The mid-questionnaires were administered after one week, and the post-questionnaires were administered at the end of the three-week study. The post-study questionnaires also included the bond factor in working alliance inventory (WAI) [12] to assess whether participants perceived a socio-emotional bond with the virtual character in the interactive journaling application.

Survey | Measurement
---|---
Pre-survey (start date) | Demographic information, Journaling Habit Questionnaire, Signature Strength Survey [10], Brief Mood Introspection Scale (BMIS), Perceived Stress Scale, Affect Balance Scale, Ryff’s Psychological Well-Being Scales
Mid-survey (after 1 week) | BMIS, Perceived Stress Scale, Affect Balance Scale, Ryff’s Psychological Well-Being Scales
Post-survey (after 3 weeks) | BMIS, Perceived Stress Scale, Affect Balance Scale, Ryff’s Psychological Well-Being Scales

Table 1. List of questionnaires used for pre-, mid- and post-tests.

Interactive Journaling Application

Eleven positive psychology interventions (Table 2) were framed as questions and programmed into the interactive journaling application. For each session with the app, three out of these eleven positive psychology interventions were used. Before and after the journaling session, participants were asked report their arousal and valence levels in numeric scores [-4, 4] (see Figure 1).

When the first set of arousal/valence scores were reported, the virtual agent appeared on the screen and greeted the participant (“Hi! Nice to see you again. Are you excited for the questions?”). The greeting utterances were varied each day within a pre-selected pool. Then, the participant pressed the button below the character to make the virtual character provide the first question and interactively journaled as much as desired. During the participant’s response, the virtual agent displayed back-channeling and listening behaviors (“Hmm”, “Uh-huh”, “I see”, “Oh”, etc.) when a verbal pause was recognized. When done answering a question, the participant went on to the next question by pressing the button again. When all three intervention questions were answered, the virtual character thanked the participant for their responses and said goodbye to exit the interaction. The self-report arousal/valence screen (Figure 1a) appeared again and the app closed after participants responded.

![Figure 1](image-url) (a) Arousal and valence assessment screen for pre- and post-interaction with the interactive journaling application, (b) virtual character at idle position, (c) virtual character animated.

<table>
<thead>
<tr>
<th>Interventions</th>
<th>Questions</th>
</tr>
</thead>
</table>
| Positive | (1) What are three good things that happened today?  
(2) Is there someone you feel gratitude to?  
(3) Tell me about one kind act you did today.  
(4) Tell me three funny things that you experienced today. |
| Neutral | (1) How did you use yours signature strengths in a new way today?  
(2) Is there anything you’d like to talk about?  
(3) How did you use your “gift” of time to someone today? It could be helping someone, sharing meal, etc. |
| Negative | (1) Was there anything that made you angry today?  
(2) Is there someone you need to forgive?  
(3) Tell me about a moment today when something bad turned into something good. |
Table 2. Lists of questions used for each positive psychology intervention in the interactive journaling application.

**Affect Detection with Affdex and SentiStrength**

We used the Affdex mobile SDK, a commercial tool marketed by Affectiva, Inc. (affectiva.com), to detect users’ emotions during smartphone usage. Affdex can detect 34 FACS (Facial Action Coding System) units [9] along with 9 emotions (joy, anger, disgust, contempt, engagement, fear, sadness, surprise, and valence) and 15 expressions (attention, brow furrow, brow raise, chin raise, eye closure, inner brow raise, lip corner depressor, lip press, lip pucker, lip suck, mouth open, nose wrinkle, smile, smirk, and upper lip raise). The valence feature was used to calculate the user’s mood, and we recorded all of the nine emotion metrics for analysis in real time with the front-facing camera on the Android device. Each of these metrics produced a score in range of [-100, 100], except for valence, which ranged [-100, 100].

While the smartphone was turned on, the application ran the facial expression capture process and captured facial emotional data from any visible face from the front-facing camera. The facial expression capture process silently started with five-minute intervals and captured data for 10 seconds. When no face was found, no affect data was recorded. While the user was using the interactive journaling application, the facial expression capture process ran continuously without stopping.

The sentiment of smart phone users’ incoming and outgoing text messages was analyzed in order to infer the users’ emotional state and to update the mood score. We used the SentiStrength tool [29,30] to evaluate the sentiment of each text message, which gave us both positive and negative scores. The positive score ranged between [1, 5] and the negative score ranged between [-5, -1]. SentiStrength was able to take account of widely used emoticons, such as :) or <3, as well as the usage of all capitalized words, for sentiment analysis. For example, the message “Cool cool. also turns out Dustin has to work :/ so I’m bringing my Friend Sam if that’s OK with you” results in a positive score of 2 and a negative score of -2, while the message “Aww, that is so sweet! Sure. I can hold onto the tickets until then. :J” produces 3 and -1. The application retrieved the content of text messages whenever there was an incoming or outgoing SMS message. The positive and negative scores for a single SMS message were added together, and then multiplied by 20 to give a score on a consistent scale with the valence score from Affdex.

Since the mood is a collective and aggregate metric of continuous affect and emotions over time, both the previous mood score and the most current valence score contributed to the daily mood score with a decay effect over time:

$$m_{t+1} = \alpha m_t + \Delta \nu(1 - \alpha)$$

where $m$ and $\nu$ are the mood and valence scores respectively. $\alpha$ is the decay rate (0.95) and $\Delta \nu$ is the elapsed time since the last mood score update time in hours. The mood score was updated whenever new affect data was available, either by capturing facial expression information through Affdex or by analyzing the sentiment of a text message via SentiStrength.

**Personalization**

The intervention selection behavior was modeled as a Markov Decision Process (MDP) [24]. The policy on the MDP model was formulated as a $Q(s, a)$ matrix, where $s$ represented the user’s mood and affect state, and $a$ represented the selection of the intervention type. The state space consisted of three dimensions: daily mood, current valence, and current engagement. The daily mood state was discretized to three values.

$$s_{\text{mood}} \in \{\text{Negative, Neutral, Positive}\} \in \{-100, 0, 100\}$$

$$s_{\text{valence}} \in \{\text{Negative, Neutral, Positive}\} \in \{-100, 0, 100\}$$

$$s_{\text{engagement}} \in \{\text{Low, High}\} \in [0, 0.91], [0.91, 100]$$

In total, the state space consisted of $3 \times 3 \times 2 = 18$ states. The action space consisted of 3 actions, $a = \{\text{Negative, Neutral, Positive}\}$. The action selection was made when the user launched the interactive journaling application for the journaling activity. Once the intervention type was chosen, three questions within the selected intervention category were randomly selected for the session. The initial policy represented an equally random distribution over three possible actions. A reinforcement-learning algorithm was implemented in order to personalize the intervention selection policy to each user. In order to achieve this, a standard SARSA (State-Action-Reward-State-Action) algorithm [25] was used. In our algorithm, the reward was calculated as a weighted sum of the valence score, engagement score, and the duration of the journaling session:

$$r = \gamma \cdot \tau_{\text{end}} - \tau_{\text{start}} + 0.3\nu_{\text{valence}} + 0.66(\nu_{\text{valence}} - \nu_{\text{valence}}_{\text{start}})$$

$\tau_{\text{start}}$ and $\tau_{\text{end}}$ are the start and end time of the journaling session in the unit of seconds. This reward function aimed to maximize engagement and valence scores, and the duration of user’s journaling activity. In order to control for the exploration and exploitation of the MDP model, a $\varepsilon$-greedy algorithm was implemented. $\varepsilon$ was set to decrease with each successive session $\varepsilon_1 = 0.75$, $\varepsilon_2 = 0.5$ and $\varepsilon_i = 0.25$ for $i >= 3$. The learning rate also decreased $\alpha_1 = 0.5$, $\alpha_2 = 0.4$, $\alpha_3 = 0.3$, $\alpha_4 = 0.2$ and $\alpha_5 = 0.1$ for $i >= 5$.

**RESULTS**

**Personalized Intervention Selection Policies**

A different intervention selection policy was developed for each participant. Figure 2 shows the Euclidean distances between participants’ final intervention selection policies. The learning algorithm personalized to each participant, and
they ended up with drastically different policies from one another. The policies did not converge after 3 weeks, but this was not surprising, since there was a large state space and relatively few learning interactions.

**Figure 2.** A distance matrix for nine participants’ final intervention selection policies after three weeks.

**Increased arousal and valence after each session**

Both arousal and valence levels increased after participants used the interactive journaling app. The mean arousal level before the journaling activity was -0.59 and increased to -0.39 after the session (Wilcoxon signed rank test, N = 137, \( p = 0.017 \)). The mean valence level before the interaction was 0.53 and increased to 0.82 after the session was over (Wilcoxon signed rank test, N = 137, \( p = 0.004 \)).

**Figure 3.** Both arousal and valence levels increased after participants used the interactive journaling app.

**A trend of increase in self-acceptance level**

Pre-, mid- and post-test responses showed no statistically significant differences in participants’ mood, perceived stress and affect balance. However, participants’ self-acceptance levels from the Ryff’s Psychological Well-Being scale [26] showed a trend of difference (Friedman’s test, \( p = 0.10 \)). High self-acceptance levels indicate that a person has gained a more positive attitude toward the self and has accepted multiple aspects of self, including good and bad qualities [26]. Pairwise comparisons on self-acceptance levels failed to show a statistically significant difference with \( p<0.05 \), but we did see a trend of increase from the mid-test to the post-test (Tukey’s test, \( p = 0.08 \), Figure 4).

**Figure 4.** Self-acceptance levels showed a trend of increase between mid-test and post-test

**Increased interaction duration over time**

We compared participants’ response length for each question in the first 1.5 weeks of the study to the second 1.5 weeks. Figure 5 shows that the responses lasted a mean of 24.62 seconds for each intervention question in the first half of the study phase, and 31.59 seconds for the second half. This increase in the length of response time was statistically significant (Mann-Whitney U test, N = 677, \( p = 0.0002 \)).

**Figure 5.** Interaction length per question increased in the second half of study.

**Intimate and personal journal entries**

Participants talked about intimate and personal matters. Although the prompts asked mainly about specific events or actions and were rather limited, participants expanded their responses to reflections and evaluations of themselves, their relationships with others, their purpose of life, and more. Below is a transcription of one journal entry made in the study. \( \text{A} \) represents the virtual agent of the app and \( \text{P} \) indicates the participant’s response.

**A:** How did you use your signature strength in a new way today?

[Transcription of journal entry]

---

They ended up with drastically different policies from one another. The policies did not converge after 3 weeks, but this was not surprising, since there was a large state space and relatively few learning interactions.
P: (clears throat) I don't remember what my signature strength is, I think it's "appreciation" ... maybe? Uhm ... what did I do today? I fell asleep during class ... lunch ... oh okay, so, yeah, last night my friend came up, uhh, to basically ... to basically just talk about, like, random stuff. But, uhm, we ended up talking for, like, an hour-and-half about ... my life, and my ... and I don't know, it got kind of deep? Like my character, or my world views and how maybe there are things that I can work on, which I agree with (clears throat). And then I talked to my roommate about some more, sort of like ... views on change, or how ... I was very, like, interesting conversations.

Uhm, and then it's something that I've been thinking about a lot and so I kind of ... this is more like the gratitude question, I guess. But I'm grateful that they took time out of their day to listen to me. I feel like all I do is complain, which is another thing I'm working on. And so it was nice that ... I really did ... I really did appreciate that they were there and that they listened to what I had to say and they made time for me and ... 

A: (Excited) Oh! Tell me anything you would like to talk about.

P: Uhm ... I guess just generally I feel like ... getting ... getting better, I guess, or ... ? Maybe closer to who I was before, and the next step is for me to let go of all these negative thoughts and feelings that I still have. I don't know why, unconsciously or subconsciously, clinging to these thoughts (clears throat). Like I feel like I have to be ... or it's not that I feel like I have to be, I feel like ... I have this weird idea that my natural state is negative. And I feel like that's not true ... because ... 

A: (Excited) Mmm!

P: I feel like things have changed ... and I want to change them back. Because I liked who I was as a person more before I sort of entered this storm of negativity. Uhm, I'm going to do my writing exercise tonight and I think since writing is a good way to sort out your thoughts and explore ... deeper what you truly believe, to organize these things that you have, I'll probably use that time to kind of figure out how I'm feeling about who I am ... and sort of the way that I've been acting lately.

A: How did you use your gift of time to someone today? It could be helping someone, sharing a meal, or anything like that.

P: Uhm, I shared a meal with the tennis team today. I kind of wanted to just dream box and come back to my room and watch T.V. and ... eat (laughter) by myself ... because sometimes I'm antisocial. But (clears throat), I ended up eating with the team and an old teammate that came back to visit us this week and it was nice ... talking to them. And I feel like the fact that I think it was nice, or maybe even just some part of me ... means that I'm moving in the right direction because like if you asked me a month ago I probably would have been insufferable and I feel like yeah it's the whole ... vicious cycle thing, like my negativity is making me act a certain way, people are reacting accordingly and then I'm just sort of like "This is terrible". But ... I don't know, that was a positive thing that happened today, I guess.

A: Thank you for sharing. It was good to hear your stories. See you!

DISCUSSION

Even though no statistically significant change in participants’ psychological wellbeing was found across the three-week study, the application had the short-term effect of increasing arousal and valence levels for its users. This result aligns with previous research on expressive writing therapy [18-20], and shows that the journaling framing of the intervention activity was effective in improving users’ affect.

Participants interacted longer in the second half of the study than in the first half. This could have resulted from the unique intervention selection policies that the reinforcement algorithm learned for each user over time, but an alternative hypothesis also exists. Another cause could be that over the course of three weeks, participants became more familiar with the activity and learned to express themselves better with the application. However, it is difficult to make a conclusive interpretation of what caused the change in interaction length because there was no control group that used the application without the personalization to compare with the treatment group. A follow-up study with multiple experimental treatment groups would be needed in order to identify the effect of the personalization algorithm on users’ engagement with the app.

The qualitative analysis of the recorded journal entries provided some evidence that participants reflected about themselves, which could have led to the trend of increase in self-acceptance levels. This was intriguing because the intervention questions in the journaling application did not directly ask participants about their opinions of the self. However, participants often described themselves (e.g., “I usually try to have positive attitudes,” “I get irritated really easily,” “I am the unhealthiest person in the world.”) while describing circumstances or events that happened to them for their journal entries.

Sometimes, these discussions of the self led participants to find new meanings and appreciation for themselves. For instance, during an informal interview at the end of the study, one participant reported that she felt answering the “three kind acts” question was difficult initially because she did not perform big charity or community service work on daily basis. However, she soon realized that even small acts, such as preparing a nice warm meal for her family or friends, could be considered a kind act, and she found meaning in her daily chores that she used to consider trivial. On the other hand, another participant confessed in one of her journal entries that she was an “insecure” person and explored the idea whether she should feel “gratitude” for the friendship that she and her friend shared. She commented that it felt “weird” to say that she is grateful that her friend likes her since a friendship is supposed to be a “mutual thing.” However, during a later journal entry, the same participant also narrated her appreciation and
gratitude toward her friends for listening to her despite the fact that “all [she] does is complaining”.

LIMITATIONS AND FUTURE WORK
We interviewed the participants at the end of the study to identify limitations of the interactive journaling application. Some participants commented that they were frustrated at the repetitiveness of the interventions. They suggested having a bigger pool of questions, and even additional open-ended questions for the activity. Many participants mentioned that they sometimes talked to the character about matters that were not strictly related to the question because they had the desire to reflect about certain events that were significant to them. Also, a few participants expressed difficulties listing “three” items for funny things or good things that happened to them. The “three funny things” had significantly lower answering rate than any other intervention questions.

Another limitation of the application was that the virtual character always talked in high-pitched and cheerful voice. Sometimes the character’s cheerful greetings, e.g., “Hi! Hope you had a wonderful day,” irritated the users when they were in a negative mood. Addressing the emotional match between the character and the user can be significant in providing a sense of empathy and emotional bond [13].

Based on participants’ feedback, we plan to expand the set of questions and to enable the character to identify the user’s emotion and match the tone of its voice as well as the interventions to the user’s current emotional state. We will also run a randomized control trial experiment with the personalized and non-personalized application in order to evaluate the efficacy of the personalization.

CONCLUSION
We developed a smartphone application that provided an interactive journaling experience for users to improve their affect and mood. The system successfully engaged users for three weeks, and users engagement with the application increased over time. We used facial expressions and SMS messages to detect users’ emotional states, and we used the collected affect data to personalize the human-agent interaction for each user. Despite some weaknesses, most users found the application helpful in reflecting about themselves and their daily lives. Our work could complement existing psychological interventions and enable people to have more frequent and easy access to interventions that improve human wellbeing.

ACKNOWLEDGEMENT
This work was funded by LG Electronics Inc.

REFERENCES

136


