Evaluating a Personalizable, Inconspicuous Vibrotactile (PIV) Breathing Pacer for In-the-Moment Affect Regulation

Pardis Miri, Emily Jusuf, Andero Uusberg, Horia Margarit, Robert Flory, Katherine Isbister, Keith Marzullo, James J. Gross
Stanford University, University of Maryland, and University of California, Santa Cruz, Intel Labs
parism@stanford.edu, ejusuf@stanford.edu, andero@stanford.edu, andero@stanford.edu,
robert.flory@intel.com, katherine.isbister@ucsc.edu, marzullo@umd.edu, gross@stanford.edu

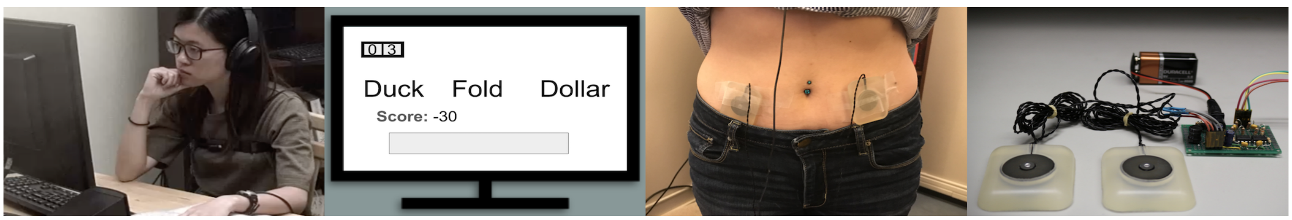

Figure 1. A participant solves compound remote associate questions under a time limit on a computer. Two tactors, attached to the abdomen, deliver personalized paced vibrations with which participants can synchronize their breathing. The answer to the question above would be "bill," a word which is related to all three words shown.

ABSTRACT
Given the prevalence and adverse impact of anxiety, there is considerable interest in using technology to regulate anxiety. Evaluating the efficacy of such technology in terms of both the average effect (the intervention success) and the heterogeneous effect (for whom and in what context the intervention was effective) is of paramount importance. In this paper, we demonstrate the efficacy of PIV, a personalized breathing pacer, in reducing anxiety in the presence of a cognitive stressor. This is the first mixed-design study of a vibrotactile affect regulation technology which accounts for individual differences and user-technology engagement in relation to the technology's efficacy in the presence of a specific stressor. Guidelines in this paper can be applied for designing and evaluating other affect regulation technologies.

Author Keywords
Haptic; Vibrotactile; Anxiety; Affect Regulation; Affect; Respiration; Slow-paced breathing; Pacer; Wearable; Linear Mixed Model; XGBoost; Shapley Values; Machine Learning;

CCS Concepts
•**Human-centered computing** → **Human computer interaction (HCI);** *Haptic devices;* User studies;

CHI '20, April 25–30, 2020, Honolulu, HI, USA.

ACM ISBN 978-1-4503-6708-0/20/04 ...$15.00.
http://dx.doi.org/10.1145/3313831.3376757

INTRODUCTION
Affect regulation refers to the things that we do to influence experience and expression of affective states [17].[1] Regulating high-arousal negative affective states is a known challenge. There is significant interest in developing technology that can help people regulate high-arousal negative affect, such as anger, anxiety, and stress, in everyday life [1–5, 7, 9, 12, 13, 15, 30, 34]. It is crucial that any affect regulation technology adds value and can eventually be effective in everyday life. Therefore, it would be useful for evaluation of affect regulation technology to understand both user-technology engagement and the heterogeneous effect (for which individuals a technology has the potential to be effective under a given stressor).

In this paper, we evaluate the efficacy of a high-fidelity prototype of PIV, a personalizable, inconspicuous vibrotactile breathing pacer developed to guide users through slow-paced breathing to reduce anxiety (see Figure 1d). In [25], we introduced PIV pacer and provided recommendations on appropriate placement, pattern, and personalization of it. [2]. In this paper, using these recommendations, we demonstrate the PIV pacer to be a successful intervention for anxiety reduction, and provide insights on who is more likely to benefit from it.

We ran a mixed-design experiment in which a treatment and control group went through two repeated stressors. Only the treatment group received intervention from PIV, which occurred during the second stressor. The analysis of our experiment had three aims. As part of open science best prac-

[1] Affect is an umbrella term for emotion, mood, and stress.
[2] Placement:abdomen; Pattern: biphasic frequency-based pattern with a less intense vibration indicating the inhalation phase

tices [6, 24], the first aim was pre-registered and the three pre-registered hypotheses can be found at [27]. The second and third aims were exploratory analyses.

1. *Efficacy of PIV in regulating affect.* Does PIV effectively regulate affect? To answer this question, we measured three affect-related dependent variables: anxiety as measured by STAI-6 scores [23], self-reported positive affect, and self-reported negative affect. For each of these dependent variables, we pre-registered the hypotheses that we would observe interaction effects between group (treatment and control) and condition (Post-stressor 1 and Post-stressor 2).[3] Our hypotheses were that the treatment group would show a decrease in STAI-6 scores, an increase in positive affect, and a decrease in negative affect compared to the control group when going from Post-stressor 1 to Post-stressor 2.

2. *Stressor and PIV-user engagement.* How does the presence of a stressor impact a user's ability to engage with PIV? We measured difficulty of PIV-user engagement at three levels: noticing vibrations, differentiating between inhalation and exhalation vibrations, and synchronizing breathing with vibrations. The stressor we used was a timed compound remote associate task [8] (see Figure 1b and the Protocol section for details). We sought to understand how the presence of this specific stressor could contribute to changes in difficulty of noticing, differentiating, and synchronizing with PIV.

3. *Predicting anxiety drop from individual differences.* What types of people have potential to benefit from PIV? We collected measures of individual difference before and within the study. Using these measures as features, we trained a machine learning model to predict drop in anxiety in the treatment group. We calculated the average Shapley values of each feature to identify the factors that contributed most to the model's prediction. Shapley values measure the contribution of a feature in predicting some label, in our case drop in anxiety [22, 32]. From the highest contributing features, we identified the types of people who have more potential to benefit from PIV.

The specific contributions of this paper are as follows:

- We report the results of a thorough, mixed-design, pre-registered experiment to evaluate the efficacy of PIV.
- We quantify the relationship between our stressor and PIV-user engagement.
- We trained a learning model on treatment group data and report the average Shapley values of features with regards to the model's predicted change in anxiety. By identifying features with the highest contribution to predicted anxiety change, we present insights on who is more likely to benefit from PIV.

Our study not only evaluates the efficacy of PIV, but also provides a thorough report on the effect of a specific stressor on difficulty of engaging with PIV, as well as who is more likely to benefit from PIV. Our findings, experimental design, and methods of analysis may be of benefit to others in the HCI community who are building and evaluating affect regulation technology with the goal of eventually having it be used in daily life. We envision a future in which a mental health clinician could choose between various technological interventions, selecting the ones that are more likely to be suitable for a patient, given knowledge of that patient.

[3] Post-stressor 1 and Post-stressor 2 are the conditions immediately following the two repeated stressors, respectively. See Figure 2 for where these conditions are in the experimental protocol.

COMPARISON WITH RELATED WORK

A list of vibrotactile technologies for affect regulation may be found in the Related Work section of the PIV publication [25]. In addition to those in the list, BoostMeUp [13] and Spire [34] are vibrotactile technologies described in more recent publications. Of the vibrotactile technologies for affect regulation, we decided to compare PIV with technologies which had been evaluated for efficacy using an experimental design comparable to ours. Specifically, we looked for an experiment that delivered a stressor to a treatment group (who received vibrotactile intervention) and a control group (who received no vibrotactile intervention). Therefore, BoostMeUp was excluded for comparison, because the experiment to evaluate it did not have a control group that received no intervention. Spire was similarly excluded for comparison, because the experiment to evaluate it was longitudinal. To the best of our knowledge, the only two vibrotactile technologies evaluated using a design comparable to ours are Doppel [7] and EmotionCheck [12].

To reduce anxiety, PIV applies slow breathing-like vibrations to the abdomen (see Figure 1c), while Doppel and EmotionCheck apply slow heartbeat-like vibrations to the wrist. All three are *involvement* technologies, i.e. technologies that aid a user in implementation of affect regulation.[4] In Doppel and EmotionCheck, affect regulation is implicit: the user engages with the technology at no more than the perception or unconscious cognition level, and they do not need to expend effort or take action. In PIV, affect regulation is explicit: the user engages with the technology at all three levels of perception, cognition, and action (that is, noticing, differentiating, and sychronizing) to intentionally implement guided slow-paced breathing.

Our study to evaluate PIV builds on the methods used to evaluate Doppel and EmotionCheck in two key ways. First, we used a mixed design in our evaluation of PIV, whereas the evaluations of Doppel and EmotionCheck used between-subjects designs. A mixed design enables greater statistical power than a between-subjects design. It also helps to address threats to internal validity, meaning that it better enables experimenters to rule out alternative explanations for observed results. Second, we collected measures of state and trait affect regulation, as well as personality, in our study. State and trait affect regulation refer to the strategies individuals use and the success

[4] Terminology is taken from Miri et al., who describe a framework for categorizing affect regulation technology [26, 29]. The extended process model of affect regulation identifies four steps in affect regulation: *identifying* the need for affect regulation, *selecting* an affect regulation strategy, *implementing* the strategy, and *monitoring* how well the implementation is going [18]. In the absence of technological support, all four steps are done by the individual. Affect regulation technology is categorized based on the step(s) in which a technology is involved.

	Doppel [7]	EmotionCheck [12]	PIV
User-technology interaction	Implicit involvement	Implicit involvement	Explicit involvement
Experimental design	Between-subjects design; N = 52 across two groups, treatment and control	Between-subjects design; N = 67 across four groups	Mixed design; N = 97 across two groups, treatment and control groups
Stressor type	Anticipation of public speaking task with deception	Anticipation and delivery of public speaking task with deception	Creative thinking task with deception
State affect measures	STAI-20, administered twice	STAI-20, administered twice	STAI-6 and positive and negative affect, administered five times
State and trait affect regulation measures	No	No	Yes
Personality trait measures	No	No	Yes
Pattern shape	Heartbeat-like vibrotactile pattern	Heartbeat-like vibrotactile pattern	Biphasic pattern with different frequencies for inhalation and exhalation phases
Pattern pace	Personalized, ranging from 0.67 Hz to 1.08 Hz	Static, 1 Hz	Personalized, ranging from 0.08 Hz to 0.15 Hz
Effect size found	d = .28	d = .38	***d* = .33**

Table 1. Comparison of three studies evaluating the efficacy of a vibrotactile technology for affect regulation.

they experience in regulating affect, both at a specific time and context during the study (state) and as an overall pattern in various contexts (trait) [16, 19]. Using these state and trait measures as well as other measures, we identified who would be more likely to benefit from PIV. Neither of the other studies reported on the relation between individual differences and the efficacy of their technology in regulating affect.

Table 1 compares in more detail the technologies and the experiments used to evaluate efficacy of PIV, Doppel, and EmotionCheck. Because our mixed design required a repeatable stressor, we had to use a different stressor type compared to the studies for Doppel and EmotionCheck. The other studies used a variation on anticipation and delivery of a public speaking task as their stressor. Since this stressor could not be repeated while maintaining participants' belief in the deception, our study used a timed compound remote associate task (CRA), which asked participants to come up with a fourth word that completes a set of three seemingly unrelated words under a time limit [8] (see Figure 1b for an example). We used the STAI-6, a brief six-item version of the State-Trait Anxiety Inventory (STAI-20), in order to be able to collect anxiety measures more frequently while minimizing participant burden. STAI-6 scores correlate highly with scores on the full-length STAI-20 ($r = .95$) [23]. In addition to differences in experimental protocol, the PIV technology differs from Doppel and EmotionCheck in the characteristics of the vibrotactile pattern used. Doppel and EmotionCheck deliver vibrations modeled after the human heartbeat. The pace of Doppel is personalized to approximately 20% lower than the user's baseline heart rate, with fixed upper and lower limits of 1.08 Hz (65 BPM) and 0.67 (40 BPM) respectively. The pace of EmotionCheck is fixed at 1 Hz (60 BPM). PIV, on the other hand, delivers vibrations modeled after human respiration. The vibrations of PIV are biphasic with a less intense vibration indicating the inhalation phase and a more intense vibration indicating the exhalation phase. The cycle is repeated at a personalized pace which was found to generally fall within the range of 0.08 to 0.15 Hz (5 to 9 breaths per minute).

Sample size determination was based on the effect size of 0.28 found in the Doppel study. We found this effect size reasonable and predicted an effect size of 0.28 or greater in our study. Using a power analysis with power of 80% and alpha of 0.05, we selected a sample size of N = 100.

METHOD AND PARTICIPANTS

Participants were recruited through university pools offering either course credit or payment. The study was deceptively advertised as an attempt to "investigate the relationship between haptics and creativity." Eligible participants had no history of cardiovascular, breathing, neurological, or psychological disorders, and restricted alcohol and caffeine consumption as well as smoking for a period of time before the on-site study. We required that participants be native English speakers because of the vocabulary-based CRA stressor task.

100 participants were randomly assigned to treatment and control groups while balancing for gender. One participant was excluded from the analysis because they were using their phone during the Stressor 2 condition. Two other participants were excluded due to procedural errors. After excluding those participants whose sessions were invalid, our sample size was N = 97 with 44 (29 female, 15 male) in the treatment group and 53 (32 female, 21 male) in the control group.

Treatment and control groups were found to be balanced across a variety of characteristics. There was no significant difference between treatment and control groups in years of education, area of education, previous training in slow-paced breathing, or whether they were currently playing a wind instrument.

Participants were fitted with two tactors taped to the abdomen according to the recommendation of the PIV derivation study.[5] They were also fitted with breathing gauges around the chest and abdomen; EDA, pulse, and temperature sensors on their

[5]For a detailed description of the components of the PIV prototype, see [25]. For the list of all during-the-study questions see the supplementary material.

non-dominant hand; and noise-cancelling headphones to block out the sound of vibrations.

The on-site portion of the study took place in the following stages: Baseline, Meditation, Personalization, Vibrotactile Breathing Practice, Stressor Task Practice, Pre-stressor 1, Stressor 1, Post-stressor 1, Pre-stressor 2, Stressor 2, and Post-stressor 2 (see Figure 2). The Pre-Stressor 1, Stressor 1, and Post-stressor 1 stages are collectively referred to as Block 1, while the Pre-stressor 2, Stressor 2, and Post-stressor 2 stages are collectively referred to as Block 2. Participants spent 60 to 90 minutes in total for the on-site portion of the study.

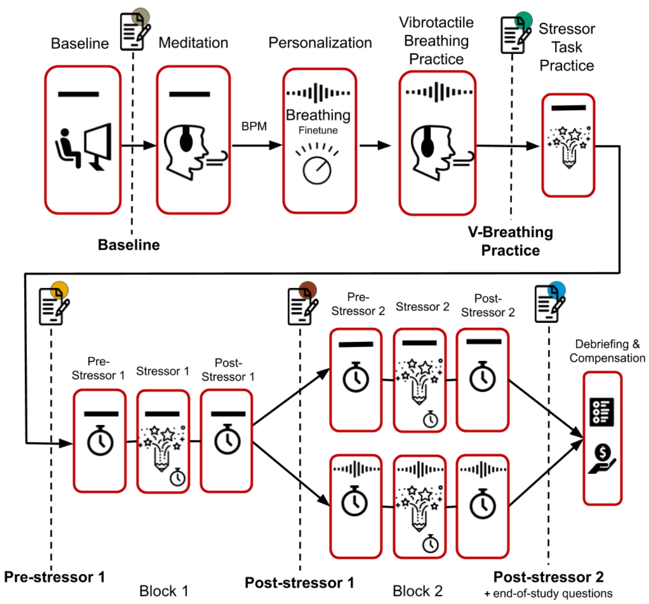

Figure 2. Procedure flowchart of the study. All participants went through the same procedure until end of Post-stressor 1. At that point, participants who were randomly assigned to the control group went through Block 2 while receiving no vibrations (indicated by a solid black line), while participants who were randomly assigned to the treatment group go through Block 2 while receiving vibrations. Self-reported measures of affect (positive and negative) as well as STAI-6 were collected at five conditions of Baseline, V-Breathing Practice, Pre-stressor 1, Post-stressor 1, and Post-stressor 2 (indicated by color-coded pen-and-paper icon).

In the Baseline stage, participants watched a 5-minute video of a pipe changing colors and were asked to count how many times the pipe turned green. The objective of this stage was to obtain baseline measurements.

In the Meditation stage, participants listened to a 5-minute audio guiding them through slow-paced abdominal breathing techniques. The objective of this stage was to train participants in preparation for breathing with the vibrations, and to estimate their slow-paced respiratory rate in breaths per minute.

The Personalization stage established an appropriate pace, frequency, and amplitude of vibrations personalized to the participant. Personalized pace was informed by the respiratory rate measured during the last two minutes of the Meditation stage. Research assistants worked with participants to arrive at appropriate personalized frequencies and amplitudes.[6]

In the Vibrotactile Breathing (V-Breathing) Practice stage, participants practiced synchronizing their breathing with the personalized vibrations for 90 seconds.

In the Stressor Task Practice stage, participants were introduced to the compound remote associate task (CRA) to be used as a stressor. Compound remote associate questions consist of three English words whose answer is a fourth word that is associated with all of them. For example, the answer to the CRA question "man/glue/star" would be "super." Participants solved three CRA questions with no time limit in order to familiarize themselves with the task.

Block 1 consisted of the first stressor and two waiting periods occurring before and after the stressor. Waiting periods, called Pre-stressor 1 and Post-stressor 1, allowed time to pass in between stressors so that anxiety from the stressor would not affect measurements taken at other stages. During the Pre-stressor waiting periods, participants waited for two minutes while seeing the text: *You are about to start a set of problems, each with a 9 second time limit. Your goal should be to solve these problems to the best of your abilities. You will be compared to your peer-group and you will receive information about your performance compared to them at the end of the session*. During the Post-stressor waiting periods, participants waited for two minutes while seeing the text: *Wait for the next part to begin automatically*.

The stressor task was four minutes long and consisted of 27 CRA questions with a time frame of nine seconds each. Running out of time on a question was equivalent to an incorrect answer. If a question was answered correctly, participants would still have to wait the full nine seconds before seeing the next question; this was done to ensure that all participants spent an equal amount of time in the stressor task. Participants were able to see the nine-second timer as well as their cumulative score. A recorded voice provided dynamic accuracy feedback by announcing "Correct" or "Incorrect." Due to the presence of sensors on the non-dominant hand, participants were required to type with only one hand, which increased the difficulty of the task.

All participants followed the same procedure from Baseline up until the end of Block 1. The structure of Block 2 was identical to that of Block 1. In addition to the previously mentioned instructions, the treatment group saw the following additional instructions during Pre-stressor 2: *Please allow your breathing to conform with the sensation of the pacer if you can, whenever you can.* Having two blocks which were otherwise identical except for the presence of vibrations in the treatment group was done to ensure internal validity: since participants' anxiety and affect in Block 2 were being compared to their anxiety and affect in Block 1, we could be more certain that any change we observed was due to the vibrations and not due to other differences between participants. During Block 2,

[6]For more information on the personalization part of the protocol, see Supplementary Materials of PIV paper, "Fine-Tuning Horizontal Pattern" section in Personalization Routine Script.

the treatment group received ongoing personalized vibrations while the control group did not.

Specific CRA questions were taken from Bowden & Jung-Beeman [8]. To ensure that the two sets of CRA questions in Stressors 1 and 2 were of similar difficulty levels, we performed K-means clustering on data from Bowden & Jung-Beeman about the number of people who correctly answered each question within different time frames. We used the clustering results to identify pairs of questions with similar difficulty levels; for each pair, one question was randomly assigned to Stressor 1 and the other to Stressor 2. More detail about our clustering and pairing process can be found in Supplementary Materials Section Stressor Design [28].

Measures

We collected state affect measures, state and trait affect regulation measures, user-technology engagement measures, and personality trait measures.

Trait affect regulation measures and personality measures were collected from an online survey which participants completed before arriving for the on-site portion of the study. The survey contained six standardized questionnaires yielding 11 variables in total. The Big Five Inventory yielded scores on the five personality factors of extraversion, agreeableness, conscientiousness, neuroticism, and openness [20]. The Anxiety Sensitivity Index yielded a score representing the extent to which an individual believes that anxiety experiences have negative implications [31]. The Body Sensations Questionnaire yielded a score representing how worried an individual feels about bodily sensations associated with autonomic arousal [11]. The Reducer-Augmenter Scale yielded a score representing an individual's preferences for the intensity of sensory input [14, 35]. The Emotion Regulation Questionnaire yielded two scores representing how much an individual uses the affect regulation strategies of cognitive reappraisal and expressive suppression [19]. The Difficulties in Emotion Regulation Scale was kept as a single score (sum of all subscales) representing the extent to which an individual struggles with affect regulation [16].

State affect measures were collected at five conditions (time points in the protocol, called conditions to distinguish from stages in the protocol). Conditions correspond to the stage of the on-site study after which they are named: Baseline, V-Breathing Practice, Pre-stressor 1, Post-stressor 1, and Post-stressor 2. Immediately after each stage, participants took the STAI-6 questionnaire and rated their positive and negative affect on a scale of 0 to 100. We followed the unipolar valence model in asking participants to describe affect in terms of positive and negative affect [21]. We gave participants the following instructions in rating positive and negative affect: *People often report their feelings as a combination of some positive and some negative feelings at once. For example, a young woman who had just eaten a chocolate bar reports a blend of joy and guilt. To report your feelings, please consider using both positive and negative feelings scales.*

User-technology engagement measures were collected at two conditions: V-Breathing Practice (no stressor) and Post-stressor 2 (stressor). User-technology engagement refers to how difficult participants found it to engage with the breathing pacer at three distinct levels of perception, cognition, and action [10]. At the perception level, we asked how difficult it was to notice the vibrations. At the cognition level, we asked how difficult it was to differentiate between inhalation and exhalation phases of the vibrations. At the action level, we asked how difficult it was to synchronize their breathing with the vibrations. All ratings were on a scale ranging from 0 to 100. Because only the treatment group received vibrations during Block 2, there is no data from the control group on engagement at the Post-stressor 2 condition.

State affect regulation measures were collected at the Post-stressor 2 condition. We asked participants to rate the extent to which they deployed each of the affect regulation strategies of reappraisal, distraction, suppression, and slow-paced breathing to lower anxiety during the study [33]. All ratings were on a scale ranging from 0 to 100. Questions were phrased in such a way so that participants did not need to be familiar with affect regulation terms. For instance, reappraisal refers to cognitive reframing of an emotional event so as to have a different emotional meaning. The statement to rate for reappraisal was "I tried to think in a way that helped me stay calm." Suppression refers to inhibition of behaviors associated with emotional response, including facial expressions, verbal utterances, and gestures. The statement to rate for suppression was: "I tried to suppress my anxiety."

In addition to collecting state affect regulation measures at Post-stressor 2, we asked supplementary questions at the end of the study to the treatment group about their opinions on the breathing pacer. On a scale of 0 to 100, we asked them to rate how much the vibrations *impacted* them; how much they wanted to *turn off* the vibrations; how much they felt the vibrations were *distracting*; how much the vibrations affected their *performance*; and what percentage of the time they were able to *synchronize* with the vibrations.

RESULTS AND DISCUSSION

Efficacy of PIV in Regulating Affect

We ran linear mixed models on STAI-6 scores and self-reported positive and negative affect. We checked for outliers before running the models and did not find any.[7]

We tested for interaction effects between condition (Post-stressor 1 and Post-stressor 2 conditions) and group (treatment and control groups). We found an interaction effect between condition and group for STAI-6 scores, but not for positive and negative affect. To characterize the interaction effect that we found for STAI-6 scores, we performed paired and independent samples t-tests.

STAI-6: We observed an interaction effect between condition and group for STAI-6 scores (intercept = 40.0, $\beta = -3.19$, p = .019, CI = $[-5.71, -.433]$) (Figure 4). The direction of the effect is consistent with our hypothesis, indicating that

[7] Models were implemented in R with the lme4 package according to the following specifications: `lmer(DV ~ condition * group + (1 | id))`.

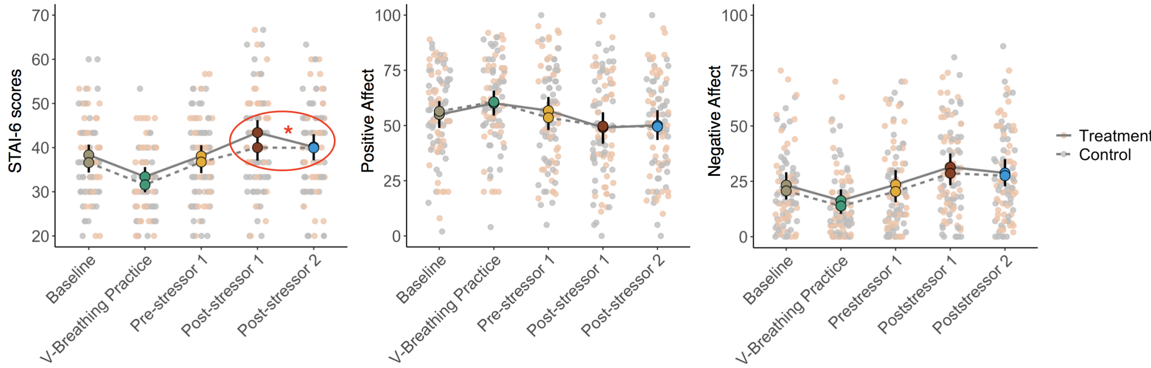

Figure 3. Left: Average STAI-6 scores during five conditions for both treatment and control groups. An interaction effect (circled) was observed between group and post-stressor 1 and post-stressor 2 conditions: the treatment group receiving vibrotactile patterns during post-stressor 2 experienced a drop in anxiety compared to the control group. Solid line indicates treatment group; dotted line indicates control group. Center, Right: Average positive affect (left) and negative affect (center) during five conditions as reported on a scale of 1 to 100 by both treatment and control groups. No interaction effect was observed between group (treatment and control) and condition (post-stressor 1 and post-stressor 2): we did not find that receiving vibrotactile patterns had an influence on either positive or negative affect.

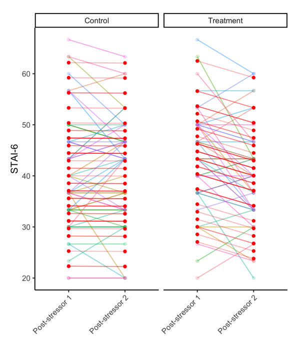

Figure 4. Linear mixed model fit to STAI-6 scores. The model fits horizontal lines for the control group and lines with negative slopes for the treatment group (shown in red), in both cases considering individualized intercepts. The model predicts a significant drop in anxiety for the treatment group going from the post-stressor 1 to post-stressor 2 condition.

the treatment group experienced a drop in anxiety from Post-stressor 1 to Post-stressor 2 compared to the control group. We calculated the effect size (Cohen's d) of the drop in anxiety for each of treatment and control groups. In the treatment group, the effect size for STAI-6 scores between Post-stressor 1 and Post-stressor 2 was moderate (mean of differences = $-3.26, d = .33$) (Figure 3, left); in the control group, the effect size was minimal (mean of differences = $-0.06, d = .05$). We performed paired t-tests to confirm that the drop in anxiety was indeed significant for the treatment group and not for the control group. As expected, STAI-6 scores between Post-stressor 1 and Post-stressor 2 were significantly different for the treatment group ($p = .004$), but not for the control group ($p = .93$). [8].

Positive and Negative Affect: We did not observe any interaction effects between condition and group for either positive or negative affect (Figure 3, center and right). Although the direction of effects for both positive and negative affect were consistent with our hypotheses, the effects were not significant. The treatment group experienced an increase in positive affect from Post-stressor 1 to Post-stressor 2, but the effect was not significant (intercept = 49.1, β = 13.1, p = .53, CI = $[-2.96, 5.34]$). Likewise, the treatment group experienced a decrease in negative affect from Post-stressor 1 to Post-stressor 2, but the effect was not significant (intercept = 31.4, $\beta = -1.56$, p = .58, CI = $[-7.28, 3.98]$). We postulate that one reason we did not see significant effects is because both the stressor and the intervention specifically targeted anxiety, whereas positive and negative affect each have many components.

Having observed an interaction effect between condition and group for STAI-6, we ensured that the interaction effect was only present between Post-stressor 1 and Post-stressor 2 conditions and not between any of the prior condition pairs. Both treatment and control groups underwent an identical procedure up until the end of Post-stressor 1, and the two groups were found to be balanced across a variety of characteristics; however, the observed mean of the treatment group at Post-stressor 1 condition (mean = 43.4) was greater than that of the control group (mean = 40.0). We therefore sought to confirm that there were no interaction effects between group and any of the prior conditions of Baseline, V-Breathing Practice, Pre-stressor 1, and Post-stressor 1. To test for the existence of interaction effects, we ran a covariance pattern model with unstructured error structure in SPSS (F(95) = .59, p = .62). The covariance pattern model was chosen instead of a linear mixed model because the linear mixed model assumes compound symmetry (equal variances and covariances across all conditions), which was inappropriate here as participants were performing different activities across the four conditions. The model did not find any interaction effects, which we interpret to suggest that there was no meaningful difference between treatment and control groups prior to the end of Post-stressor 1. We therefore conclude that the difference in the group means at Post-stressor 1 is due to noise.

[8]The found effect is not post-hoc significant. However, we bootstrapped the effect CI calculation which suggests the existence of an effect.

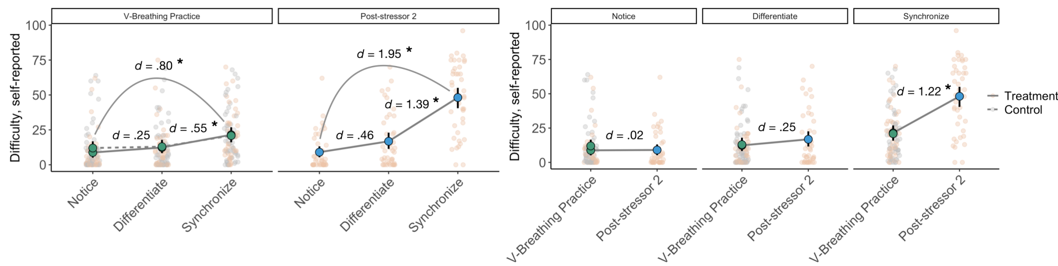

Figure 5. Left and Right: Average user-technology engagement difficulty at three levels of perception, cognition, and action in the presence and absence of a stressor, with effect sizes displayed. Left: Synchronization with vibrations is more difficult than both noticing and differentiating vibrations. This is true both without a stressor (at V-Breathing Practice) and with a stressor (at Post-stressor 2). Right: When a stressor is introduced, synchronizing becomes significantly more difficult than both noticing and differentiating. Note that there is data for both treatment and control groups in the absent stressor condition (at V-Breathing Practice), but only for the treatment group in the stressor condition (at Post-stressor 2).

Stressor and PIV-User Engagement

We sought to quantify the level of difficulty engaging with PIV at three levels of perception, cognition, and action, both in the presence and in the absence of a CRA stressor. We ran a linear mixed model on participants' ratings of difficulty of engaging with the technology at three levels: noticing vibrations, differentiating between inhalation and exhalation vibrations, and synchronizing breathing with vibrations. Both treatment and control groups rated their engagement difficulty at the V-Breathing Practice condition, while only the treatment group rated their engagement difficulty at Post-Stressor 2.[9] We tested for interaction effects between condition (V-Breathing Practice and Post-Stressor 2 conditions) and engagement type (noticing, differentiating, synchronizing).

We observed a main effect of synchronization across V-Breathing Practice and Post-stressor 2 conditions (β = 10.86, CI = $[6.58, 19.58], p$ = .0002). In both conditions, synchronizing breathing with vibrations was significantly more difficult than noticing vibrations ($d = .80, p = .0004$ at V-Breathing Practice; $d = 1.95, p < .0001$ at Post-stressor 2). In addition, in both conditions, synchronizing breathing with vibrations was significantly more difficult than differentiating vibrations ($d = .55, p = .01$ at V-Breathing Practice; $d = 1.39, p < .0001$ at Post-stressor 2) (Figure 5, left). These results suggest that affect regulation strategies which involve synchronization with a slow-paced breathing pacer are difficult regardless of whether or not a CRA stressor is present.

In addition to the main effect of synchronization, we observed an interaction effect between engagement type and condition (Figure 5, right). When a stressor is introduced, synchronizing breathing becomes significantly more difficult compared to both noticing and differentiating vibrations (β = 26.43, CI = $[17.22, 36.01], p < .0001$ for the change between notice and synchronize; β = 20.52, CI = $[11.25, 30.13], p < .0001$ for the change between differentiate and synchronize). Figure 5, right, shows effect sizes (Cohen's d) of changes in engagement difficulty from V-Breathing Practice to Post-stressor 2. These results suggest that synchronizing becomes more difficult in the presence of a CRA stressor, while other types of engagement do not necessarily become more difficult in the presence of a CRA stressor.

[9]The model was implemented in R according to the following specifications: `lmer(engagement-difficulty ~ engagement-type * condition + (1 | id))`.

Predicting Anxiety Drop from Individual Differences

We sought to identify factors that could explain the drop in anxiety observed in the treatment group. We pooled 26 individual difference measures collected either before or within the study, and trained a learning model on a dataset consisting of treatment group observations with those 26 features. For each feature, average Shapley values contain information about that feature's contribution to the model's predicted change in anxiety from Post-stressor 1 to Post-stressor 2. For the features with the highest contributions, we conducted further tests to determine whether the predicted drop in anxiety was due to use of PIV, or whether people with certain characteristics were already predisposed to see a drop in anxiety regardless of use of PIV. Because we had data from both treatment and control groups, we were able to make inferences about PIV's efficacy with high internal validity. That is, we had the data to determine whether the results we saw were due to intervention from PIV or due to other systematic differences between individuals. We found a main effect of within-study reappraisal usage, suggesting that individuals who used more reappraisal during the study saw a drop in anxiety regardless of use of PIV. We also found interaction effects suggesting that certain categories of people were more likely to see a drop in anxiety when using PIV. People with low Openness on the Big Five Inventory, people with low scores on the Difficulties in Emotion Regulation scale, people with high Extraversion on the Big Five Inventory, and people with low within-study distraction usage were all more likely to benefit from PIV than their counterparts on the other end of the spectrum for each feature.

Model Selection Procedure

The 26 individual difference features in the dataset used to train the model consisted of all state affect measures, state and trait affect regulation measures, user-technology engagement measures, and personality trait measures. Though the 11 measures collected before the study are conceptually different from the 15 measures taken within the study, we pooled them together for the purpose of training the model because we wanted to compare each feature's contribution to predicted anxiety change.

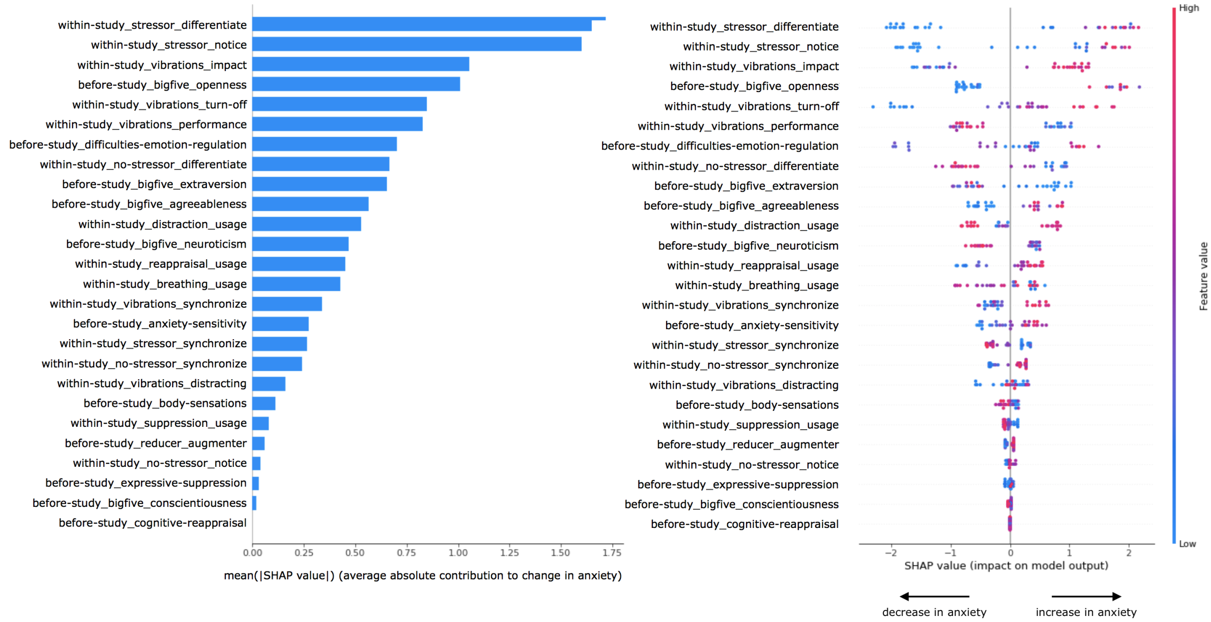

Figure 6. Left: SHAP feature importance measured as the mean absolute Shapley values. The rating of how much a participant desired to turn PIV vibrations off was the most important feature, changing the predicted anxiety level on average by 1.42 points. Right: SHAP summary plot showing the importance and the effect of features. Low numbers of willingness to turn-off vibrations contribute to anxiety drop, and large numbers to increase in anxiety.

We then trained a small number of XGBoost regression models while optimizing their hyperparameters and reported their averaged performance in terms of goodness of fit on a reserved set of testing data. Then we reported a randomly selected model, which had 219 trees with maximum depth of 11 and averaged performance of -.30 of goodness of fit. More detail about model training can be found in Supplementary Materials Section Model Selection [28].

To interpret the decision criteria for the selected model, we extracted and visualized Shapley values using the SHAP package in Python [22]. Shapley values represent, for each given observation, the contribution of a particular feature to the learning model's prediction. Each individual observation is therefore associated with 26 Shapley values, one for each feature. Shapley values can also be averaged over a single feature to obtain a measure of that feature's importance to the model prediction. We averaged Shapley values separately for each of training and test data and for each of the 26 features. In training data, Shapley values indicate what oddities in the training data motivated the model to form its decision criteria. In test data, Shapley values indicate how a model would use its structure of reasoning on unseen data. We chose to focus here on reporting Shapley values of the training data. This is because we were not able to show that the feature rankings based on Shapley values in test and training data are significantly different (Wilcoxon signed rank test: V = 155, p = .85). Figure 6, left, shows the magnitude of average Shapley values of training data for each of the 26 features. Features are sorted vertically with the greatest magnitude of contribution at the top.

Figure 6, right, uses the results of the above feature ordering to display more detailed information about each feature's influence on the prediction for each individual. For each feature, 32 Shapley values, each corresponding to an individual in the treatment group, are distributed horizontally. The color represents the value of the feature. For instance, blue points in the row "within-study_vibrations_turn-of" represent participants who did not have much desire to turn off vibrations (low rating out of 100), while pink points represent participants who wanted to turn off vibrations. The more negative the Shapley value, the more a feature contributes to drop in anxiety for that participant. We can therefore draw conclusions about the direction of contribution. For example, the model predicts that individuals who did not have much desire to turn off vibrations would see a drop in anxiety, whereas individuals who wanted to turn off vibrations would see an increase in anxiety.[10] Below, we present and analyze the most important contributing features according to average Shapley values.

Overview of Top Contributing Features

We were most interested in the contribution of before-study features, since a clinician could collect these measures to determine a patient's a priori likelihood of benefiting from a technology before the patient had interacted with the technology. As seen in Figure 6, left, the before-study feature with the greatest contribution was Openness (8% contribution). For this feature, we tested for the main effect to determine whether a drop in anxiety was observed in both treatment and control groups. If there was a main effect, it would mean that people with that characteristic already tended to see a drop in anxiety regardless of intervention from PIV. We did not find a main

[10] Points which do not conform to this pattern, such as the few individuals who had a positive Shapley value despite not having much desire to turn off vibrations (blue), result from interaction between the feature in question and some other feature.

effect for this feature. We therefore proceeded to test for an interaction effect, which would indicate an interaction between an individual difference feature and the PIV intervention to facilitate drop in anxiety. To represent the individual difference features categorically, we binned Openness into low and high, with a cutoff of 33. We chose this cutoff value from Figure 6, right, by approximating the value at which points transitioned from blue to red. We found an interaction effect between bin (low and high) and group (treatment and control). Individuals with Openness score lower than 33 saw a drop in anxiety in the treatment group, but not in the control group (intercept = 30.07, $\beta = -7.2$, p = 0.003, CI = $[-12.13, -2.176]$). See Figure 7, left. The results suggest that individuals with low as opposed to high Openness are more likely to benefit from PIV. We suspect that because openness is (slightly) positively correlated with reappraisal use, it is possible that people low on openness are less able to use other affect regulation strategies, and so benefit more from using PIV.

We were also interested in state affect regulation within-study features. These features measure how much a participant used different affect regulation strategies while stressed. Reappraisal and Distraction were the two with the highest contribution to model prediction according to average Shapley values (3% and 4% respectively). Following the same reasoning as above, we tested first for the main effect of reappraisal and distraction usage. We found a main effect of Reappraisal but not of Distraction. This suggests that individuals who used more reappraisal during the study, compared to individuals who used less, saw a drop in anxiety regardless of intervention from PIV (intercept = 44.56, $\beta = -.12$, p = .04, CI = $[-.24, -.007]$). We binned Distraction usage into low and high categories with a cutoff of 21, and tested for an interaction effect between bin (low and high) and group (treatment and control). We found an interaction effect indicating that individuals in the low Distraction category had a drop in anxiety in the treatment group, but not in the control group (intercept = 36, $\beta = -6.93$, p = .002, CI = $[-10.96, -2.78]$). However, this result is inconclusive: we saw a significant difference in anxiety from Pre-stressor 1 to Post-stressor 1 between treatment and control groups. We have no explanation for this observation, because both groups went through the same protocol procedure through Post-stressor 1. See Figure 7, right.

We also looked at the other within-study measures besides state affect regulation. The top three within-study measures that had the greatest contribution to model prediction were difficulty differentiating and noticing vibrations in the presence of a stressor (contributions of 12% and 10% respectively) and desire to turn off vibrations during the study (contribution of 8%). All three features relate to engagement with PIV: the first and third relate to engagement at the cognitive level, and the second to engagement at the perceptual level. For all three features, the pattern of Shapley values in Figure 6, right, shows an association between less difficulty and drop in anxiety. The results suggest that in general, a less effortful user experience with PIV is conducive to seeing a drop in anxiety.

The total contribution of individual difference measures was 31% and the total contribution of within-study measures was 69%. Because both individual difference and within-study measures had a significant contribution to model prediction, it was important to collect both types of measures.

DESIGN IMPLICATIONS

In this section, we address the following three questions. Given that we found PIV to be effective in reducing anxiety, can we further improve the efficacy of PIV? Given that we found a main effect of Reappraisal, does this mean that people skilled in reappraisal are less likely to benefit from PIV? Lastly, given that a less effortful user experience was conducive to anxiety reduction from PIV, does this mean that implicit involvement technologies are more effective than explicit involvement technologies?

Can we further improve the efficacy of PIV?

Yes: we think that efficacy can be further improved by adding to the personalization procedure. In the above analysis, we identified the top three within-study factors that differed between individuals and that contributed the most to anxiety reduction from PIV. These factors—difficulty differentiating vibrations in the presence of a stressor, difficulty noticing vibrations in the presence of a stressor, and desire to turn off vibrations during the stressor—have to do with experiencing vibrations in the presence of a stressor. Therefore, we believe that the personalization procedure should include exposure to a controlled stressor. Recall that we found synchronization with vibrations to be significantly more difficult when a CRA stressor was present than when it was not. Following the current last step of the personalization procedure in which users synchronize breathing with vibrations, there could be an additional step in which users synchronize breathing with vibrations in the presence of a stressor. We also recommend following this additional step with a question about how much on a scale of 0 to 100 the user wanted to turn off the vibrations during the personalization procedure. Since we found that desire to turn off vibrations was a good predictor of not seeing an anxiety drop with PIV, any answer greater than 15 (suggested by Shapley values cutoff) would be good reason to go back and try different personalization parameters.

Are people skilled in reappraisal less likely to benefit from PIV?

To answer this question, we binned reappraisal into low and high categories, with a cutoff of 71 based on Shapley values (see Figure 6). We found an interaction effect indicating that individuals in the low reappraisal category had a drop in anxiety in the treatment group, but not in the control group (intercept = 41.8, $\beta = -4.09$, p = .01, CI = $[-7.42, -0.92]$). We saw no anxiety change in the high reappraisal category, however. See Figure 7, right. These results together suggest that people less skilled in reappraisal are more likely to benefit from PIV, as compared to those who are more skilled. We also observed that the average anxiety of those less skilled (mean = 35.2) in reappraisal is statistically lower than those less skilled (mean = 41.71, p = .006). This suggests that while PIV can be helpful for those who are less proficient in reappraisal, being skilled in reappraisal is more effective overall. We think it is promising that PIV can help those who are not innately

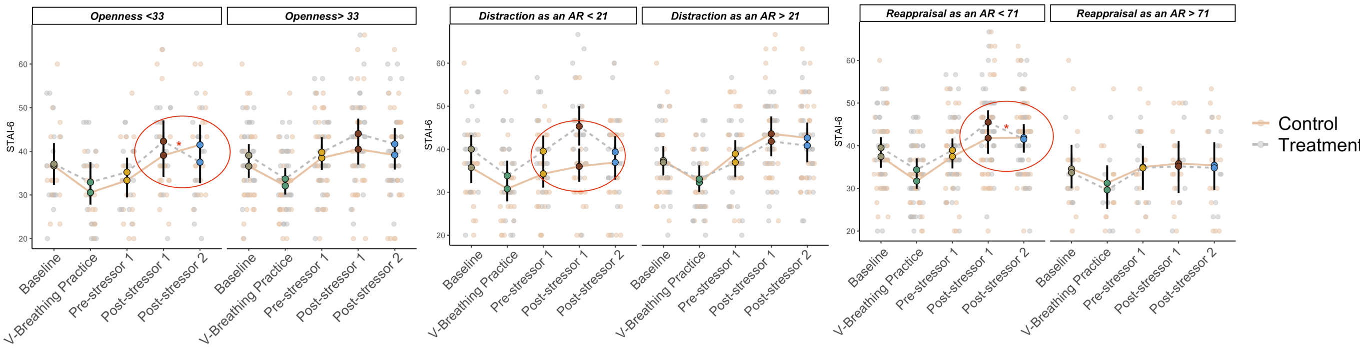

Figure 7. Left: Average STAI-6 scores. An interaction effect (circled) was observed between group and Post-stressor 1 and Post-stressor 2 for Openness score < 33 but not for Openness score > 33. Center: An interaction effect was observed between Pre- and Post-stressor 1 as well as between Post-stressor 1 and 2 which make the results inconclusive. Right: The same as left but for Reappraisal scores cutoff at 71.

skilled in affect regulation. The question remains whether prolonged use of PIV by people who are less proficient in affect regulation can improve their proficiency. We recommend a follow-up longitudinal study to determine if this is the case.

We recommend that others in the HCI community similarly evaluate existing affect regulation technologies with regards to individual differences. Collecting these measures will provide insight on how to improve a technology's efficacy. It will also ensure that affect regulatory technology is able to benefit the populations for which it is meant—namely those who struggle with affect regulation and seek technology to facilitate it.

Are implicit involvement technologies more effective than explicit involvement technologies?

What makes PIV an example of explicit involvement technology is that the user intentionally and effortfully synchronizes their breathing with vibrations. Yet we found that synchronization was more difficult than engagement at both perceptual and cognitive levels, and more difficult with a stressor present than without. In addition, the top three within-study measures contributing to predicted anxiety drop all suggested that less effortful interaction with PIV was conducive to anxiety reduction. Taken together, these results raise the question of whether implicit involvement technology, which by definition requires no effort, is a better choice than explicit involvement.

We respond by pointing out that it is unclear from our results whether less effortful interaction with PIV led to a drop in anxiety by having an effect at the perceptual and cognitive levels (implicit regulation) or by facilitating synchronization (explicit regulation). We recommend a follow-up study to determine the mechanism of the effect from less effortful interaction by looking at physiological measures of breathing. By determining whether and when users were actually breathing in sync with their vibrations, we should be able to determine whether they derived an implicit or explicit regulation benefit from PIV. Until then, we cannot answer the question of whether implicit or explicit involvement technologies are more effective.

CONCLUSION AND FUTURE WORK

In this paper, we presented a thorough evaluation of PIV by reporting on its efficacy in reducing anxiety, the relation between stressor and PIV-user engagement, and the contribution of various factors to predicted change in anxiety. We ran a mixed-design experiment in which treatment and control groups each went through two repeated CRA stressors. We found an interaction effect between condition and group, indicating that intervention from PIV helped to reduce anxiety in the treatment group but not in the control group with effect size of $d = .33$ [11]. We also quantified the relation between the CRA stressor and PIV-user engagement; we found that synchronizing breathing with PIV was more difficult than other forms of engagement and more difficult with a stressor present than without. We suggest that in future studies, personalization of PIV should involve synchronization in the presence of a stressor. Finally, we trained a machine learning model and calculated the average Shapley values of 26 individual difference features in order to predict who might benefit most from PIV. Further analysis on these features suggested that individuals who are low on Openness and within-study reappraisal usage are more likely to benefit from PIV. This is a first step towards allowing potential users of PIV to predict, based on their traits and experience with the technology, whether they would benefit from PIV or not. We hope other researchers interested in building affect regulation support technologies will find our results, as well as our methodological and analytic strategies, useful in future work. In particular, we hope our work will encourage researchers to report on specific stressors, user-technology engagement, and individual differences when evaluating affect regulation technology. In seeking to make affect regulation technology that will be effective outside the lab, it is important to understand not only the efficacy of a technology, but the conditions that affect its performance. [12]

[11] We observed a lack of correlation between psychological arousal inferred from skin conductance and the STAI-6. When ignoring the breathing patterns, we did not observe an interaction effect between group and condition for the psychological arousal inferred from skin conductance, due to presence of a cognitive stressor which plays a role in how much participants were able to follow the patterns. See the supplementary material for more details.

[12] All code, anonymized datasets, and learning model discussed in this paper is available on github https://github.com/paris007/PIV_plus_plus.

REFERENCES

[1] *Touchpoint*. https://thetouchpointsolution.com. Accessed 8/31/2019.

[2] 2018. *Lief Patch*. https://www.youtube.com/watch?v=aclzen-GGTE. Accessed 8/31/2019.

[3] 2018. *Lief Therapeutics*. https://getlief.com. Accessed 8/31/2019.

[4] 2018. *Spire Stone*. https://spire.io/pages/stone. Accessed 8/31/2019.

[5] 2018. *Vitali Sports Bra*. https://vitaliwear.com. Accessed 8/31/2019.

[6] Chris Allen and David MA Mehler. 2019. Open Science challenges, benefits and tips in early career and beyond. *PLoS biology* 17, 5 (2019), e3000246.

[7] Ruben T Azevedo, Nell Bennett, Andreas Bilicki, Jack Hooper, Fotini Markopoulou, and Manos Tsakiris. 2017. The calming effect of a new wearable device during the anticipation of public speech. *Scientific Reports* 7 (2017). DOI:http://dx.doi.org/10.1038/s41598-017-02274-2

[8] Edward M Bowden and Mark Jung-Beeman. 2003. Normative data for 144 compound remote associate problems. *Behavior Research Methods, Instruments, & Computers* 35, 4 (2003), 634–639.

[9] breathe App. https://support.apple.com/en-us/HT206999. Accessed 6/7/2019.

[10] Charles S Carver and Michael F Scheier. 1982. Control theory: A useful conceptual framework for personality–social, clinical, and health psychology. *Psychological bulletin* 92, 1 (1982), 111.

[11] Dianne L Chambless, G Craig Caputo, Priscilla Bright, and Richard Gallagher. 1984. Assessment of fear of fear in agoraphobics: the body sensations questionnaire and the agoraphobic cognitions questionnaire. *Journal of consulting and clinical psychology* 52, 6 (1984), 1090.

[12] Jean Costa, Alexander T Adams, Malte F Jung, François Guimbetiere, and Tanzeem Choudhury. 2016. EmotionCheck: leveraging bodily signals and false feedback to regulate our emotions. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 758–769. DOI: http://dx.doi.org/10.1145/2971648.2971752

[13] Jean Costa, François Guimbretière, Malte F Jung, and Tanzeem Choudhury. 2019. BoostMeUp: Improving Cognitive Performance in the Moment by Unobtrusively Regulating Emotions with a Smartwatch. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3, 2 (2019), 40.

[14] Steven Dragutinovich. 1987. Australian factorial confirmation of Vando's Reducer—Augmenter Scale. *Personality and individual differences* 8, 4 (1987), 489–497.

[15] Asma Ghandeharioun and Rosalind Picard. 2017. BrightBeat: effortlessly influencing breathing for cultivating calmness and focus. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. ACM, 1624–1631.

[16] Kim L Gratz and Lizabeth Roemer. 2004. Multidimensional assessment of emotion regulation and dysregulation: Development, factor structure, and initial validation of the difficulties in emotion regulation scale. *Journal of psychopathology and behavioral assessment* 26, 1 (2004), 41–54.

[17] James J Gross. 2002. Emotion regulation: Affective, cognitive, and social consequences. *Psychophysiology* 39, 3 (2002), 281–291.

[18] James J Gross. 2015. The extended process model of emotion regulation: Elaborations, applications, and future directions. *Psychological Inquiry* 26, 1 (2015), 130–137.

[19] James J Gross and Oliver P John. 2003. Individual differences in two emotion regulation processes: implications for affect, relationships, and well-being. *Journal of personality and social psychology* 85, 2 (2003), 348.

[20] Oliver P John, Sanjay Srivastava, and others. 1999. The Big Five trait taxonomy: History, measurement, and theoretical perspectives. *Handbook of personality: Theory and research* 2, 1999 (1999), 102–138.

[21] Peter Kuppens, Francis Tuerlinckx, James A Russell, and Lisa Feldman Barrett. 2013. The relation between valence and arousal in subjective experience. *Psychological Bulletin* 139, 4 (2013), 917.

[22] Scott M Lundberg, Gabriel G Erion, and Su-In Lee. 2018. Consistent individualized feature attribution for tree ensembles. *arXiv preprint arXiv:1802.03888* (2018).

[23] Theresa M Marteau and Hilary Bekker. 1992. The development of a six-item short-form of the state scale of the Spielberger State—Trait Anxiety Inventory (STAI). *British Journal of Clinical Psychology* 31, 3 (1992), 301–306.

[24] Erin C McKiernan, PE Bourne, CT Brown, S Buck, A Kenall, J Lin, D McDougall, BA Nosek, K Ram, CK Soderberg, and others. 2016. How open science helps researchers succeed, ELife 5. 16800 (2016). DOI: http://dx.doi.org//10.7554/elife

[25] Pardis Miri, Robert Flory, Andero Uusberg, Heather Culbertson, Richard Harvey, Agata Keman, Erik Peper, and Keith Gross, J. James Marzullo. 2019. PIV: Placement, pattern, and personalization of an inconspicuous vibrotactile breathing pacer. (2019). DOI: http://dx.doi.org//10.1145/3365107

[26] Pardis Miri, Andero Uusberg, Heather Culbertson, Robert Flory, Helen Uusberg, James J. Gross, Keith Marzullo, and Katherine Isbister. 2018. Emotion Regulation in the Wild: Introducing WEHAB System Architecture. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, LBW021. DOI: http://dx.doi.org/10.1145/3170427.3188495

[27] Miri Pardis, Gross James J., Marzullo Keith, Uusberg Andero, and Jusuf Emily. 2019a. *Pre-registered Hypothesis*. (2019). https://osf.io/epn3h.

[28] Miri Pardis, Gross James J., Marzullo Keith, Uusberg Andero, and Jusuf Emily. 2019b. *Supplemental Material*. (2019). https://osf.io/rbk2h/.

[29] Heather Culbertson Robert Flory Helen Uusberg James J. Gross Keith Marzullo Katherine Isbister Pardis Miri, Andero Uusberg. 2018. *Emotion Regulation in the Wild: The WEHAB Approach*. Technical Report. University of California, Santa Cruz. https://www.soe.ucsc.edu/sites/default/files/technical-reports/UCSC-SOE-18-04.pdf.

[30] Pablo E Paredes, Yijun Zhou, Nur Al-Huda Hamdan, Stephanie Balters, Elizabeth Murnane, Wendy Ju, and James A Landay. 2018. Just breathe: In-car interventions for guided slow breathing. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 1 (2018), 28. DOI: http://dx.doi.org/10.1145/3191760

[31] Steven Reiss, Rolf A Peterson, David M Gursky, and Richard J McNally. 1986. Anxiety sensitivity, anxiety frequency and the prediction of fearfulness. *Behaviour research and therapy* 24, 1 (1986), 1–8.

[32] Lloyd S Shapley. 1953. A value for n-person games. *Contributions to the Theory of Games* 2, 28 (1953), 307–317.

[33] Gal Sheppes, Gaurav Suri, and James J Gross. 2015. Emotion regulation and psychopathology. *Annual review of clinical psychology* 11 (2015), 379–405. DOI:http://dx.doi.org/10.1146/annurev-clinpsy-032814-112739

[34] Eric N Smith, Erik Santoro, Neema Moraveji, Michael Susi, and Alia J Crum. 2019. Integrating wearables in stress management interventions: Promising evidence from a randomized trial. *International Journal of Stress Management* (2019). DOI: http://dx.doi.org/10.1037/t02889-000

[35] Alan Vando. 1974. The development of the RA scale: A paper-and-pencil measure of pain tolerance. *Personality and Social Psychology Bulletin* (1974). DOI: http://dx.doi.org/10.1177/014616727400100110