The Mobile Stress Meter: A New Way to Measure Stress Using Images

Sophia Haim
Dartmouth College
sophia.haim@dartmouth.edu

Lorie Loeb
Dartmouth College
lorie@dartmouth.edu

Rui Wang
Dartmouth College
rui.wang@cs.dartmouth.edu

Xia Zhou
Dartmouth College
xia@cs.dartmouth.edu

Sarah E. Lord
Dartmouth College
sarah.e.lord@dartmouth.edu

Andrew T. Campbell
Dartmouth College
campbell@cs.dartmouth.edu

Abstract
Stress is an important aspect of human psychology. Many measures of stress have been developed over the years, but different issues exist for most of these measures. This study introduces the mobile stress meter, a new way to measure stress in which a user simply selects an image that best captures his or her stress level. Such a tool allows researchers to quickly measure stress in real time and in natural environments less onerous for the user. Our results show that the mobile stress meter is a valid measure of stress. Users find it easy and actually enjoyable to use. Our results also show that the mobile stress meter is strongly correlated with the Perceived Stress Scale, a validated multi-item stress scale. The correlation has an r-value of 0.5559 and a p-value of less than 0.001. We conclude that the mobile stress meter can be applied effectively in research experiments and advance the research on stress.

Introduction
Stress has been defined in many ways. Possibly the simplest definition of stress is a physical or psychological state of strain [1]. A more complete definition describes stress as the physiological or psychological response to internal or external stressors [14], where a stressor is a thought or event that conveys threat,
harm, or loss [7].

The normal stress response is actually beneficial: it promotes adaptive changes when faced with an acute threat. These changes can include increased vigilance, attention and arousal, as well as a shutdown of feeding behaviors [3]. The stress response is designed to ensure that enough energy is supplied to the organs needed to assure survival. Stress responses can also assist cognitive performance levels and enhance emotional learning and memory [7]. All of these adaptations allow humans to more readily face a threat. This suggests that stress responses can enhance performance, characterizing stress as a positive effect. In addition to this biological response, individuals can react to stress with behaviors that might be positive, negative, or neutral [6].

Biological responses to stress, though initially positive, can start to have negative effects if sustained [3]. Stress has been shown to have negative effects on health [6]. Over time, prolonged stress responses can damage tissue, weaken the immune system, cause heart damage, and lead to a range of other health problems. In addition, stress can lead to harmful behaviors, such as smoking, excessive drinking, drug abuse, and loss of appetite [1, 7]. Stress has also been associated with depression and anxiety disorders [7].

Finding ways to effectively measure stress is an important area of research, as it is the first step in recognizing and potentially reducing stress, therefore minimizing the adverse mental and physical effects of stress. There are several existing methods to measure stress, ranging from analyzing cortisol levels in saliva [6], administering multi-item questionnaires [6], to recognizing stress levels in speech [8]. While many of these methods are effective, they are invasive, expensive, annoying for users, somewhat unreliable as a measure of stress, or susceptible to recall bias. To advance the research on measuring stress, we seek a tool that can reliably measure stress, is easy to administer and use, and can be administered in a subject's natural environment, e.g., the tool can be administered on mobile phones and thus is quick and easy to use on the go.

We present a tool called the mobile stress meter, which seeks to fulfill all the above objectives. The mobile stress meter was inspired by PAM, the Photographic Affect Meter [11], but is designed to measure stress instead of affect. We follow a very similar methodology to what was used in PAM in designing our new measure of stress. The mobile stress meter is designed to be administered on smartphones, which have become very common, making it very easy for subjects to use in their natural environment at any time. Users are shown a grid of sixteen images, each of which is associated with a particular level of stress. Individuals are asked to touch the image that best represents their current stress level, and are given the option of requesting more images if they would like more options (the full image set consists of 48 images).

We carry out two studies to test the mobile stress meter. The first study asks users to choose an image and then take the Perceived Stress Scale (PSS), which has been validated in several samples and is considered the “global measure of stress” [2, 6]. The goal of this study is to determine if the mobile stress meter is a valid measure of stress, by comparing
the resulting mobile stress meter scores to the PSS scores. In the second study, 19 Dartmouth students use the mobile stress meter to record their stress level at least three times a day for seven days, starting from the second-to-last day of classes through the second-to-last day of finals. The goal of this study is to find out whether the mobile stress meter is effective and unobtrusive in actual use. These two studies lead us to conclude that the mobile stress meter is a reliable measure of stress that can easily and unobtrusively measure stress in a subject's natural environment at any time.

These results imply that a simple smartphone app can use the mobile stress meter to effectively measure stress. This is important because it will allow researchers to assess stress levels at key moments and in context, without being unpleasant to the user. This creates opportunities for a range of studies in the subject's natural environment. As smartphones have become commonplace, such studies would not require costly tools in the stress measurement process. The mobile stress meter opens up a whole new range of possibilities for stress-associated research.

**Related Work**

In an effort to make assessment of stress less invasive and burdensome for the subject, several recent studies have attempted to find ways to measure stress using smartphones and wearables. One study [12] used a wrist sensor (for accelerometer and skin conductance information), the mobile phone usage (call, SMS, location, whether the screen was on or off), and surveys (on stress, mood, sleep, tiredness, general health, alcohol intake, caffeine intake, and electronics usage) to find physiological and behavioral markers for stress. These results showed that both mobile phone usage and wearable sensor data included features related to the stress level. In particular, fewer or shorter sent SMS and certain screen on/off patterns were related to higher reported stress levels [12].

In another study [10], the authors designed a wearable sensor to unobtrusively measure electrodermal activity (EDA)/GSR over a long term. The study found evidence that the distal forearm, where the sensor was placed, is a viable alternative to the traditional encumbering palmar sites for EDA measurement. This new compact, low-cost, unobtrusive sensor offers the ability to perform comfortable, long-term, in-situ assessments of GSR [10]. Although there will still be the issue of environmental and organismal factors having an influence on GSR measurement, this new technology eliminated many of the burdensome aspects of assessing EDA/GSR. This product evolved into the Q sensor.

StressSense [8] uses smartphones to unobtrusively recognize stress from human voice. StressSense is tested on multiple individuals in several environments to make sure it can adapt to different scenarios. It provides an easy and non-intrusive way to measure stress in real-life situations. AutoSense is a wearable intended to unobtrusively and robustly measure stress in the field [5]. In a study in both the lab and the field, with over 20 subjects, AutoSense was shown to provide 90% accuracy as a model of stress.

Ecological Momentary Assessment (EMA) seeks to eliminate the recall bias and lab environment issues presented in the traditional self-report surveys. EMA involves repeated, often random, samples of subjects'
experiences and behaviors in real time and in subjects’ natural environments. Because subjects will have to complete the same assessment many times, it is important to consider that any small annoyance can become extremely irritating when encountered over and over again [13].

Mobile Stress Meter Design
We design the mobile stress meter as a stress EMA, which asks the user to select a picture that best describes his/her current stress level. The EMA methodology has been widely used to collect users’ in-situ psychological states such as affect and stress [9, 13]. EMA asks for user’s active response, which can be obtrusive so that the response compliance would drop in longitudinal studies. We believe that we can mitigate this problem by changing the form of EMA questions. During the exit interview of the StudentLife project [15], the participants reported that PAM was the least onerous EMA question among other text EMAs because image EMAs are more interesting to answer. We therefore adopt similar design in the mobile stress meter.

The mobile stress meter is designed as a 4x4 grid with 16 images as shown in Figure 2. Each cell in the grid represents a particular stress level, which we refer to as the mobile stress meter score. The least stressful images are located in the bottom left corner of the grid, while the most stressful images are located in the top right corner of the grid. Each cell/score is associated with three images (so there are a total of 48 images in the mobile stress meter). Each time the app is opened, one of these three images is randomly selected for each slot. The user is instructed to select the image that best captures how stressed they feel right now. If none of the images currently displayed seems to capture the user’s stress level, he/she can click the “More Images” button, located below the image grid as shown in Figure 1. This button will reload the grid with a new set of the mobile stress meter images. When a user chooses an image, the app pulls up a larger version of this image. The user then has the option to “Cancel” or “Submit.” If, upon seeing the larger image, the user decides it does not capture his/her stress level, the user can press “Cancel” and choose a new image. Otherwise, the user presses “Submit,” and the score, along with a timestamp, is registered. The user receives push notifications at different times throughout the day to select a mobile stress meter image.

Choosing Images. Choosing images for the mobile stress meter is nontrivial. Each picture should represent a similar stress level for most users. We therefore choose the stress image in a crowdsourcing manner: users determine what stress level each image represents, and we then validate the objective score assigned to each image by comparing image selection with PSS scores.

We use Creative Commons-licensed images from Flickr, an online photo-sharing service, as the source images. We need a set of search terms that are associated with stress to retrieve the images from Flickr. Therefore, we survey a set of undergraduate students at Dartmouth College to find out which images students associated with different levels of stress and calm. The survey consist of three questions: (1) What images do you associate with feelings of stress, (2) What images do you associate with feelings of calm, and (3) What images do you associate with feelings...
in between stress and calm. We survey a total of 28 students, and come up with 200 search terms. Some examples of the search terms we use are “beach” and “cup of tea” as calm search terms, “people yelling,” “piles of books,” and “face in hands” as stress search terms, and “to do list” and “running” as “in between” search terms. We pull images from the Flickr Creative Commons site using these search terms to put together an initial set of 624 images.

We find the candidate pictures for each stress level in two steps. We first narrow down the image set to get a smaller and more relevant stress image set. Then we use the Perceived Stress Scale to help us determine the stress level that each image represents.

To narrow down this image set and create the subset of images with which students would be most likely to identify, we create an early version of the mobile stress meter, which is a webpage with 104 image slots, with 6 images assigned to each slot. When the webpage loads, each slot is randomly filled with one of the six images designated for that slot. Images are roughly assigned to slots based on whether the search term used to find the image that has been classified by surveyed students as representing stress, calm, or something in between. “Calm” images are placed on the left, “stressful” images on the right, and “in between” ones in the middle. Users are instructed to select the image that best captures how stressed they feel right now. They are encouraged to scroll all the way down and look at all the images before making a selection. The user can click on a “More Photos” button to reload all the slots with a new random set of images. If the user keeps clicking the “More Photos” button, it will not show repeats until all six photos have been shown for each slot. This webpage is sent out to undergraduate students at Dartmouth and received 817 responses. Based on the results, we choose the images that have been clicked on most often to narrow our image set down to 96 images.

We then arrange the photos in a 4x4 grid. Each grid has 6 assigned images. These images are arranged from the most “calm” in the bottom left, to the most “stressful” in the top right, the same arrangement as the final version of the mobile stress meter. We send this webpage to undergraduate students at Dartmouth College and received 535 responses. We also collect the students’ responses to the Perceived Stress Scale survey as the groundtruth. We change the time scale of each PSS prompt from “In the past month” to “Recently” to capture the student’s most recent stress level.

**Evaluation**

We conduct two studies using the mobile stress meter to establish its validity and to determine if it can be used effectively to frequently measure stress in subjects’ natural environments. The first study involves comparing the mobile stress meter to the Perceived Stress Scale to ascertain its validity. We choose the PSS as our comparison measure because it is the most commonly employed assessment of stress [6] and it was validated for two samples of college students [2]. Because the entire process of designing and testing the mobile stress meter is done with college students, we want to compare it with a questionnaire that has been validated for a similar group. In the second study, a group of undergraduate students are asked to use an iPhone implementation of the
mobile stress meter for a week. They are instructed to select the mobile stress meter image that best represents their stress level three times a day for seven days. This allows us to assess the effectiveness of the mobile stress meter in an actual experimental setting.

**Study 1: Validation of the mobile stress meter with PSS**

The first study is designed to validate the mobile stress meter by comparing the mobile stress meter scores to PSS scores. For this study, we create a webpage with the mobile stress meter assessment. Users are instructed to select the image that best captures how stressed they feel right now. After selecting an image, the user is taken to a web-based version of PSS. The time scale of each PSS prompt is changed from “In the past month” to “Recently” to better gauge current stress levels.

The webpage is sent out to Dartmouth undergraduate students and 388 individuals participated. The mobile stress meter score ranges from 1 to 16. The PSS score ranges from 0 to 37. The mean of the mobile stress meter score is 9.41 with a standard deviation of 4.24, while the mean of the PSS Score is 19.14 with a standard deviation of 7.25.

We use Pierson correlation to determine the relations between the mobile stress meter scores and PSS scores. The mobile stress meter scores are strongly correlated with PSS scores with $r = 0.5559, p < 0.001$.

**Study 2: Use of the mobile stress meter in an Experimental Setting**

The goal of this study is to ensure that the mobile stress meter can be easily and unobtrusively completed several times a day for an extended period of time. This provides more evidence to validate the mobile stress meter as a measure of stress.

This study involves 16 undergraduate students (8 male, 8 female). Subjects are recruited via emails, in which the study is briefly described, and students are offered compensation for participating in the study. Because the mobile stress meter is initially designed for iPhones, students have to have an iPhone to participate. All the study participants are undergraduates at Dartmouth College. The student participants are between the age of 20 and 26, and the majority are 22. 12 of the students select White/Caucasian as their race, two select Hispanic or Latino, one select Black or African American, and one select Asian or Pacific Islander. The students are from a range of majors, including Economics, Computer Science, Government, Music, Engineering, English, and Environmental Studies.

Each subject downloads the mobile stress meter app onto their phone, and is instructed to complete the mobile stress meter assessment each time they receive a push notification. Subjects are told not to be too concerned if they missed a notification, but to try to complete the assessment at least three times a day. The study is administered over the seven-day period starting from the second-to-last day of classes and ending on the second-to-last day of finals. Across all participants, the mobile stress meter assessment is completed a total of 480 times over the course of the seven-day study.

**Table 1:** Results from regressing the mobile stress meter scores on academic characteristics.
A couple of times during the course of the study, subjects are asked to complete a simple multiple-choice assessment of their recent stress levels. The questions ask subjects how stressed they feel at a particular time of day, and the answer choices are: (a) not stressed at all, (b) slightly stressed, (c) stressed, (d) very stressed, or (e) extremely stressed. Although these questions have not been officially validated, they are intended to just provide a rough evaluation on whether the mobile stress meter score at particular times matches up with the subject’s simple verbal interpretation of how stressed they feel.

At the end of the study, subjects are asked to complete an exit survey. The exit survey includes questions about the subject’s experience in completing the mobile stress meter assessment several times a day for a week. We also collect the general attitude of subjects towards the study, which can be used to determine how users feel about completing the mobile stress meter regularly.

In addition, in the exit survey we include academic characteristics questions designed to learn more information about how stressed the subjects may have been during the week of the study. We ask students how many classes they are taking, when they finished classes, how many exams they had, when each exam was, how stressed they were for each exam, and whether they were stressed about graduating (for those graduating at the end of the term). We also ask a few questions to gauge each subject’s general stress level, including whether they easily become stressed, whether this finals period was particularly stressful compared to others, and whether they are usually relaxed. We expect that the answers to each of these questions could be related to the mobile stress meter scores. For example, we would expect that having more classes and more exams would increase the mobile stress meter scores, while being done with classes and usually being relaxed would decrease the mobile stress meter scores. Our exit interview also includes the PSS questionnaire. We expect that the mobile stress meter scores for the week would be positively correlated with PSS scores.

**Correlations of the mobile stress meter scores and Multiple-Choice or PSS Scores.** In order to test the relationships between the mobile stress meter scores and the multiple-choice scores, we run a correlation on these two variables. We get a correlation of $r = 0.374, p < 0.001$, which indicates a moderate positive relationship. We also run a correlation on the mobile stress meter scores and PSS scores, and get a correlation of $r = 0.314, p < 0.001$. This again indicates a strong significance and a moderate positive relationship between the two variables. These results further indicate
the validity of the mobile stress meter as a measure of stress, and show that users are able to represent their stress levels accurately using the mobile stress meter assessment.

Regressing the mobile stress meter scores on academic characteristics. To examine which academic characteristics influence the mobile stress meter score, we regress the mobile stress meter score on each variable. The results are shown in Table 1. Our results are very similar to our expectations. Taking more classes tend to increase the mobile stress meter scores, with a coefficient of 0.478 and a p-value of 0.001. Being done with classes has a negative coefficient, as predicted, but is not significant. Having more exams left, having an exam coming up in the next 24 hours, and being more stressed about an upcoming exam all have positive and significant coefficients, confirming our expectation. Having more exams left has a coefficient of 1.022 and a p-value of less than 0.001, having an exam in the next 24 hours has a coefficient of 1.401 and a p-value of 0.003, and being more stressed about an upcoming exam has a coefficient of 0.446 and a p-value of less than 0.001. Being easily stressed and being stressed about graduating both have positive coefficients, but are not significant. Having a particularly stressful finals period tend to increase the mobile stress meter scores as well, with a coefficient of 1.068 and a p-value of less than 0.001. Being usually relaxed tend to decrease the mobile stress meter scores, as we predict, with a coefficient of -0.453 and a p-value of 0.031. Finally, indicating a higher level of stress in the multiple choice questions and in the PSS questionnaire are both associated with higher mobile stress meter scores. The level of stress indicated in the multiple choice questions has a coefficient of 1.712 and a p-value of less than 0.001, indicating that marking one higher level of stress on the multiple-choice questions is associated with a 1.712 point increase in the mobile stress meter score. This positive, significant coefficient confirms our expectation that the mobile stress meter scores are higher during times of day when students verbally assess their stress level as higher. The coefficient for PSS score is 0.238, with a p-value of less than 0.001. Although some of these coefficients are not significant, all of them align with our expectations.

The mobile stress meter user experience. We use this study to assess how subjects feel about using the mobile stress meter as a measure of stress. We ask users to indicate their level of agreement/disagreement with several statements about the app in the exit survey. Overall, the responses show that users find the app easy to use, not burdensome, and actually enjoyable. Subjects are comfortable using the app and generally have a positive experience.

Each statement and the responses to the statements are shown in Table 2. Lower average rating indicates that the subjects agree more on the question. Most subjects respond more than the required three times a day, indicating that they probably enjoy performing the task. In discussing EMA, we mention that it is important to make sure that tasks that have to be completed repeatedly are not annoying for the user, as such tasks can become extremely annoying when performed over and over [13]. The mobile stress meter definitely follows this suggestion, since users generally like completing the task, and do not find it annoying. These reactions to the use of the
Table 2: The mobile stress meter user experience responses.

<table>
<thead>
<tr>
<th>Answer Options</th>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Neutral</th>
<th>Agree</th>
<th>Strongly Agree</th>
<th>Rating Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>I found completing the mobile stress meter assessment burdensome</td>
<td>12</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1.25</td>
</tr>
<tr>
<td>I enjoyed completing the mobile stress meter assessment.</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>7</td>
<td>5</td>
<td>4.06</td>
</tr>
<tr>
<td>I felt comfortable using the app.</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>13</td>
<td>4.81</td>
</tr>
<tr>
<td>I found it easy to learn how to use the app.</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>14</td>
<td>4.88</td>
</tr>
<tr>
<td>I found the app difficult to use.</td>
<td>14</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1.13</td>
</tr>
<tr>
<td>My overall impression of using the app was positive.</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>8</td>
<td>7</td>
<td>4.38</td>
</tr>
<tr>
<td>My overall impression of using the app was negative.</td>
<td>12</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1.31</td>
</tr>
</tbody>
</table>

mobile stress meter three times a day for a week indicate that the mobile stress meter can be used to frequently assess stress in experimental contexts without creating an extra burden to the subjects.

Discussion

Our results clearly show that the mobile stress meter is a valid and effective measure of stress. We have successfully designed a measure of stress that is easy to administer using smartphones, is not burdensome or annoying for users, and can effectively produce a valid real-time measure of stress in natural environments.

One concern about self-assessment measures of stress that we have not addressed is the concern that such measures cannot be adapted to other cultural contexts [6]. In fact, this study is specifically designed and tested with college students. That being said, the methodology used in this paper will not be difficult to replicate in another cultural context. Because the entire design process begins with surveying a subsample of students to generate search terms, it is straightforward to carry out the same design process in another environment. One would simply request search terms in a similar way, this time from a subsample of another cultural group. By starting the design process over in this way, the set of images could be specifically adapted to a new cultural context, and then verified in that context. The main concept of measuring stress with images would remain, as all types of cultures are likely to have some association between certain images and feelings of stress and calm. The app would simply have to be altered slightly to show a different set of images.

In several cases, single-item assessments have poor psychometric properties [6]. This does not mean that single-item assessments are invalid; in some cases they have been proven to be a valid measure of stress able to replace longer scales [4]. Furthermore, we prove the validity of the mobile stress meter, a single-item assessment, in this study. That being said, despite our positive results regarding the mobile stress meter, it should not replace PSS. There are several scenarios in which PSS would be the appropriate
measures. As with PAM, the mobile stress meter is intended specifically for situations where researchers would like to quickly measure real-time stress levels in different natural environments. It should be used to capture variability in stress levels around daily events, but the PSS questionnaire can still be used at the end of the study to assess general levels of stress throughout the experiment. In addition, the mobile stress meter can be integrated into health-tracking apps, as an additional measure of the user's health.

Conclusion

Our study showed that the mobile stress meter is a valid measure of stress. The mobile stress meter scores strongly correlate with PSS scores, demonstrating that the mobile stress meter is an effective measure of stress. In addition, the mobile stress meter was tested in an experiment in which students completed the assessment on their smartphones several times a day for a week. This experiment further confirmed the validity of the mobile stress meter, and illustrated that the mobile stress meter is an unobtrusive tool to measure stress levels in real time.

References