Towards Accurate Non-Intrusive Recollection of Stress Levels Using Mobile Sensing and Contextual Recall

Tauhidur Rahman[1], Mi Zhang[1], Stephen Voida[2], and Tanzeem Choudhury[1]

[1] *Department of Information Science, Cornell University*
[2] *Department of Human-Centered Computing, Indiana University at IUPUI*

ABSTRACT

Existing user input sampling methods used for stress-related psychotherapy are plagued with problems of low recall accuracy and high intrusiveness. In this paper, we propose a contextual recall-based self-report method using mobile sensing technology that captures contextual cues including location, activity, and environmental acoustics to aid accurate recollection of stress levels. We conducted a controlled study with 36 participants. Our experimental results suggest that contextual recall outperforms recall-based self-report method by both increasing the recall accuracy and minimizing intrusiveness to participants at the same time. Moreover, we quantified the contribution of each individual contextual cue in recollecting stress levels. We found that although participants perceived all of the contextual cues to be useful, in reality, not all the contextual cues are weighted equally during the recollection process.

Categories and Subject Descriptors

J.3 [**Computer Applications**]: Life and Medical Sciences-Health

General Terms

Human Factors, Design, Experimentation

Keywords

Mobile health, stress sensing, contextual recall, EMA

1. INTRODUCTION

Chronic stress is a major health concern and a growing problem in modern society. While stress can help people excel under pressure, medical research has shown that long-term exposure to high levels of stress can lead to various physiological and psychological disorders like hypertension, obesity and depression [3]. In the U.S. alone, stress is a public health problem, with 76% of citizens experiencing stress-related physiological symptoms and 75% facing psychological health issues [1]. Understanding how to detect and manage stress in everyday life is still an open research question. Accurately assessing perceived stress levels and identifying sources of stress is the first step in addressing this challenge.

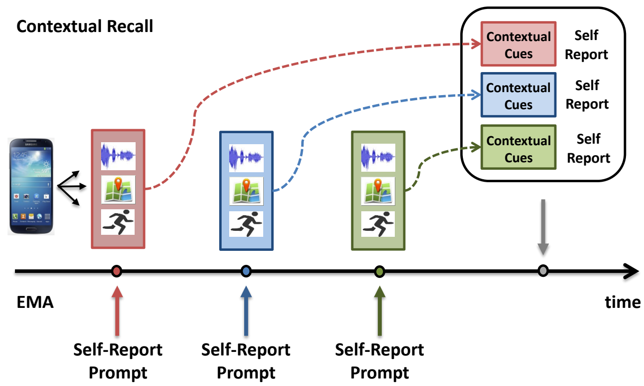

Figure 1: Comparison between EMA and the contextual recall-based self-report method

In recent years, the pervasive computing community has demonstrated some success in developing mobile and wearable sensing technologies to measure physiological responses to stress [5, 10, 9]. While these physiological measures are related to stress, the connection between these measures and *perceived stress* (how people appraise stressful situations) is tenuous [11]. Therefore, eliciting input from people is still necessary to accurately assess their perceived stress levels.

In psychotherapy of stress, both Ecological Momentary Assessment (EMA) and recall-based self-report have been adopted as gold standards for eliciting data from patients about their experienced stress and associated symptoms [14]. However, both of these methods have shortcomings. As an *in situ* user input sampling method, EMA could potentially provide the most reliable and ecologically valid inputs by minimizing the recency effect. But its high intrusiveness has been identified as leading to low compliance and high attrition in many clinical scenarios [12]. This intrusiveness is clearly undesirable for individuals suffering from stress-related problems. On the other hand, recall-based self-report minimizes intrusiveness since it asks patients to recall their past stress levels at a time of their own choosing. However, it is plagued with retrospection and rationalization biases, both of which degrade recall accuracy [14].

To address the deficiencies of EMA and recall-based self-report methods, we present a contextual recall-based self-report method that utilizes contextual cues to support accurate recall of stress levels. Figure 1 provides an illustrative comparison between EMA and our contextual recall-based self-report method. In EMA, an individual is prompted several times during the day to record his/her perceived stress level. In the contextual recall-based self-report method, contextual cues are captured using smartphone-based sensors at the moment of interest. Unlike EMA, contextual recall asks an individual to self-report their stress levels at a time of the individual's choosing, but with the help of the collected contextual cues. The only difference between recall and contextual recall-based self-report is that for recall, an individual is asked to self-report stress levels without any contextual cues.

Kahneman et al. proposed the Day Reconstruction Method (DRM), which guided participants to retrospectively tell qualitative stories about events of their choosing using a fixed set of prompts/scaffolding [8]. Although our contextual recall method is also retrospective, it uses contextual cues to guide users in recalling their stress level at a specific moment during the day. Unlike existing work that studied contextual recall for life-logging [6] and web browsing/search [7], the novel contribution of our work is the exploration of contextual recall as a means of recalling stress levels.

We make two primary contributions in this work: (i) we provide the community a clear comparison of recall accuracy for three types of user input sampling methods: EMA, recall, and contextual recall-based self-report for recording perceived stress levels; and (ii) we quantify the contribution of each individual context cue in helping participants to recall their stress levels. Our experiment was specifically designed to achieve these two goals.

2. EXPERIMENTAL DESIGN

2.1 Participants

We conducted our study on a university campus. We recruited healthy undergraduate and graduate students from a variety of different departments. In total, 36 students (20 females) between the ages of 18 and 23 participated.

2.2 Stress Induction

Since our goal was to examine how contextual recall assists in the accurate recollection of stress levels, we needed to design an experiment that would predictably induce a variety of stress levels against which the participants' self-reports could be measured. We developed an arithmetic problem solving computer game based on the empirically validated Montreal Imaging Stress Task (MIST) [4]. Our game contains twenty arithmetic problems, each at three difficulty levels: Level 1 (easy), Level 2 (medium), and Level 3 (hard) to induce different levels of stress. The game is timed (15 seconds per arithmetic problem), with the amount of time left shown at the top of the user interface. If a participant fails to solve the problem within the time limit, the game emits a loud beep and displays the message "Time Out" in red text before loading the next problem.

2.3 Selection of Contextual Cues

The contextual cues we are interested in exploring include people's location, physical activity, and the acoustic signature of their surroundings. We selected these cues because they have been identified to be critical for triggering episodic memories of past events in prior research [6, 2]. However, the role of these cues in improving stress level recall accuracy is not clear. Notice that although visual cues (e.g., images and video clips) have been identified as an important trigger of memories of past events, we intentionally avoided the usage of visual cues because of the associated privacy concerns with automatically capturing these data [6].

2.4 Capturing Contextual Cues

We used smartphones (Google Nexus 4 with Android OS) to capture contextual cues. Specifically, we used the phone's GPS sensor to collect location information, the accelerometer sensor to track physical activities, and the microphone to record the acoustic signature of the user's surroundings. It is worthwhile to note that all of these contextual cues can be collected by smartphones automatically, without any intervention from the user. Thus, mobile sensing technology significantly reduces the burden to users and serves as the foundation for contextual recall-based self-report.

2.5 Empirical Study Design

To compare EMA, recall, and contextual recall-based self-report methods and understand the role of each individual contextual cue on improving recall accuracy, we devised a between-subjects user study with six conditions. We randomly divided our 36 participants into six groups of equal size (six participants in each group) labeled as A, B, C, D, E, and F, with each group participating in a single condition. It should be noted that we used a separate group (group A) to collect the stress levels using EMA to avoid any memory priming issues. Table 1 summarizes these groups.

Group ID	Condition	Contextual Cue(s)	No. of Participants
A	EMA	-	6
B	Recall	-	6
C	Contextual Recall (with all cue)	Location, Activity, Acoustics Info	6
D	Contextual Recall	Location	6
E	Contextual Recall	Activity	6
F	Contextual Recall	Acoustic Info	6

Table 1: The group information of between-subjects user study

The study consisted of two stages: a data-collection stage and a self-reporting stage. Before the study began, participants in all six groups were asked to relax for 15 minutes to ensure that every participant began in a non-stressed mental state, thus creating a consistent baseline.

2.5.1 Data-Collection Stage

The data-collection procedure was the same for each participant across all six conditions. The purpose of this stage was to simulate real-life experiences of stress by inducing different levels of stress in different locations, activities, and acoustic surroundings. The data collection process consisted of three sessions, one after the other. In each session, the participant was escorted to one of three locations on campus with unique noise characteristics (a classroom, the library, or a restaurant) and asked to perform a certain activity (walking, sitting, or standing) while playing the stress-inducing computer game at a single difficulty level (easy, medium, or hard). After three sessions, each participant had played the game at three difficulty levels while performing three different activities at three different places. The order of locations, difficulty levels of the computer game, and activities was randomly assigned for each participant.

2.5.2 Self-Reporting Stage

For each data-collection session, each participant was asked to report his/her stress level on a 5-point scale (1 = not at all, 5 = very much) based on an existing, empirically validated self-report study of perceived stress [13]. However, the self-reporting procedure differed for each group. Participants in group A were asked to report their stress levels using a smartphone immediately following each of the three game-playing sessions, an approach closely resembling EMA. For other groups, an email questionnaire was sent to participants 24 hours after they completed the data collection stage. The questionnaire asked each participant to report his/her stress levels for each of the three sessions, optionally including various contextual cues to create different recall and contextual recall scenarios. The questionnaire also asked participants to rate the usefulness of the contextual cues, when provided. Participants in group B were asked to recall their stress levels without any contextual cues, while participants in groups C, D, E, and F were asked to recall their stress levels for all three sessions with different types of contextual cues embedded in the questionnaire (see Table 1).

3. RESULTS

3.1 Comparing Self-Report Modalities

We first examined whether the arithmetic problem-solving computer game that we built correctly induces different levels of stress. We used the stress levels reported from group A as the ground truth since EMA potentially results in the most reliable user inputs. A one-way ANOVA was conducted to compare the effect of math problem level on reported stress levels from participants in group A; the results are shown in figure 2. The statistical analysis reveals a significant main effect of math problem level on the reported stress level ($F_{2,15} = 58.81, p < .0001$). Moreover, post hoc comparisons using the Tukey HSD test show that as the math problem gets more difficult, the reported stress level increases. Together, these findings assure us that our arithmetic problem-solving computer game indeed induces distinct levels of stress that are significantly and perceivably different from one another. The randomized study parameters, including the order that the math problems were solved, the location of the test, and the physical activity undertaken during the game showed no statistically significant effects on the stress levels reported by participants. Based on these results, we ran the same statistical analyses on the stress levels reported by participants in groups B and C. We found that the main effect of math problem level on the reported stress levels is not significant for group B ($F_{2,15} = 1.50, p = .255 > .05$) but is significant for group C ($F_{2,15} = 23.59, p < .0001$). Therefore, when provided with contextual cues, people are better able to recall different levels of stress associated with different life events.

Next, we compared the reported stress levels among groups A, B and C. We ran a two-way ANOVA to compare the effects of math problem level and condition (EMA, recall and contextual recall with all cues) on reported stress levels. This analysis revealed a significant main effect of condition ($F_{2,45} = 7.30, p = .002 < .05, \eta_p^2 = .245$) on the reported stress levels. Specifically, it shows that the reported stress levels from group A are significantly higher than those from group B ($p < .05$) but are not significantly different from group C. In other words, we found that people using

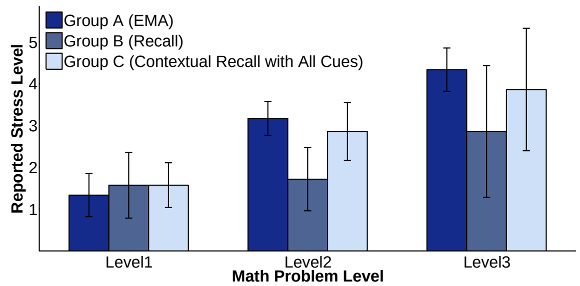

Figure 2: Comparison of stress levels across three math problem levels among EMA, recall and contextual recall with all cues

the recall-based self-report method typically underestimated their stress levels. In contrast, when assisted by contextual cues, participants retrospectively reported stress levels that were much more similar to those reported with EMA.[1]

3.2 Contribution of Different Contextual Cues

To quantify the contribution of each contextual cue as an aid for recollecting stress levels, we analyzed the reported stress levels from groups C, D, E, and F and compared them with group A, acting as the baseline. Specifically, we computed the absolute differences between the reported stress levels from groups C, D, E, and F and the mean values of the reported stress levels from group A across the three math problem levels, respectively. This absolute difference is used as the metric to evaluate the contribution of each contextual cue. The results are plotted in figure 3. Again, we ran a two-way ANOVA comparing the effect of math problem level and type of contextual cue type on the absolute differences. Our results reveal a significant main effect of contextual cue ($F_{3,54} = 4.03, p = .011 < .05$). More interestingly, post-hoc comparisons using the Tukey HSD test revealed that the difference between group E (activity) and Group A (EMA baseline) is significantly ($p < .05$) larger than the differences between groups C (all cues), D (location), and F (audio signature) and the baseline, while the differences among these groups are not significantly different from one another. These results suggest that location and the acoustic signature of the environment may be more useful cues than sensed physical activity for helping participants to accurately recall their stress levels.

Lastly, we examined the participants' ratings of their perceived usefulness of the three contextual cues (location, activity and acoustic environment) on a 5-point scale (1 = not useful, 5 = very useful). We found that all three contextual cues were perceived to be highly useful (average usefulness ratings of location, activity and acoustic environment are 4.4, 3.5 and 3.4, respectively). A one-way ANOVA reveals no significant differences among the usefulness ratings of these three contextual cues ($F_{2,18} = 2.43, p = .116 > .05$).

[1]Levene's test reveals unequal variances among stress ratings when the data from all three groups are considered together ($F_{8,45} = 3.81, p = .002 < .05$), somewhat weakening the statistical power of the reported two-way ANOVA. However, pairwise tests show that the unequal variance emerges only when including ratings from group B (recall); Levene's test confirms homoscedacity of the data drawn solely from groups A (EMA) and C (contextual recall). The combination of the results from Levene's tests and the ANOVA suggest that the ratings data from the recall condition exhibit different variance than those from the other two conditions.

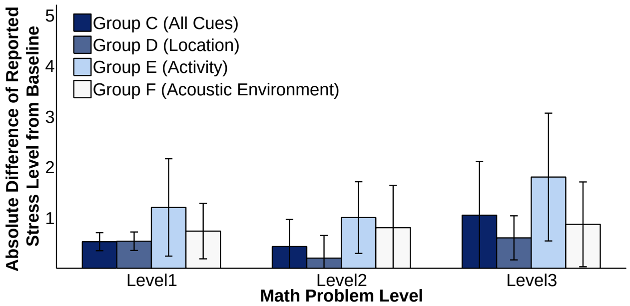

Figure 3: Absolute differences between stress levels reported by groups C (all cues), D (location), E (activity) and F (acoustic environment) and the mean levels reported by group A (EMA).

4. DISCUSSION AND CONCLUSION

We showed that contextual information captured by smartphones helped participants recollect their stress levels more accurately than when there was no contextual information provided. We also showed that contextual information can help in achieving a similar level of accuracy as EMA but without requiring *in situ* interruption. These findings indicate that contextual recall-based self-report method can remedy the shortcomings of both recall and EMA by achieving high recall accuracy and minimizing intrusiveness. EMA and recall-based self-report are widely used for recording users' stress levels; our study demonstrates that contextual recall-based self-report could be a promising new alternative.

Although participants perceived all contextual cues to be useful, in reality, not all cues were equally effective in assisting recollection. We found that participants using location or acoustic information cues were able to recall their stress levels with very high accuracy. This indicates that these cues contain important information to help participants recall stress levels. On the other hand, we found that physical activity information is not as useful. One potential explanation for this finding is that the activities this study explored were rather low-level activities (walking, sitting, standing); such low-level activities may not be effective triggers for remembering stressful events. In the future, we would like to explore whether high-level, semantically meaningful activities (e.g., taking an exam) could achieve better performance.

As a final note, we discuss some limitations of our study. First, our experiment used an artificial (but well-validated) math problem-based stress induction in a controlled setting. The stress we induced might differ from real-life stresses. We would like to run a longitudinal study in real-world scenarios to explore contextual recall of stress. Second, we used a 24-hour time window to see issues related to recall-based self-report and study the difference between recall and contextual recall. We did not explore other recollection delays because we were able to observe differences after 24 hours, and because optimizing the recall delay was not our primary focus. However, we would like to explore the optimal time duration for contextual recall-based self-report in future work.

In this paper, we explored contextual recall as a non-intrusive self-report method to help people recollect stress levels associated with stressful events. We showed that with the help of automatically recorded cues, contextual recall outperformed the recall-based self-report method and achieved similar performance to EMA in terms of recall accuracy. Moreover, we found that location and acoustic information from the surroundings are more useful contextual cues than physical activity for assisting recollection of stress levels.

5. ACKNOWLEDGMENTS

This work was partially supported by Intel Science and Technology Center for Pervasive computing, Swiss NSF Sinergia grant and NSF IIS#1202141.

6. REFERENCES

[1] American Psychological Association. Stress in America. http://www.apa.org/pubs/info/reports/2007-stress.doc.

[2] S. Carter and J. Mankoff. When participants do the capturing: The role of media in diary studies. In *Proc. CHI*, pages 899–908. ACM, 2005.

[3] S. Cohen, R. Kessler, and L. Gordon. *Measuring stress: A guide for health and social scientists*. Oxford University Press, New York, 1997.

[4] K. Dedovic et al. The Montreal Imaging Stress Task: Using functional imaging to investigate the effects of perceiving and processing psychosocial stress in the human brain. *J. Psychiatry and Neuroscience*, 30(5):319–325, 2005.

[5] E. Ertin et al. AutoSense: Unobtrusively wearable sensor suite for inferring the onset, causality, and consequences of stress in the field. In *Proc. SenSys*, pages 274–287. ACM, 2011.

[6] R. Gouveia and E. Karapanos. Footprint tracker: Supporting diary studies with lifelogging. In *Proc. CHI*, pages 2921–2930. ACM, 2013.

[7] J. Hailpern et al. YouPivot: Improving recall with contextual search. In *Proc. CHI*, pages 1521–1530. ACM, 2011.

[8] D. Kahneman, A. Krueger, D. Schkade, N. Schwarz, and A. Stone. A survey method for characterizing daily life experience: The day reconstruction method. *Science*, 306(5702):1776–1780, 2004.

[9] H. Lu et al. StressSense: Detecting stress in unconstrained acoustic environments using smartphones. In *Proc. UbiComp*, pages 351–360. ACM, 2012.

[10] K. Plarre et al. Continuous inference of psychological stress from sensory measurements collected in the natural environment. In *Proc. IPSN*, pages 97–108. ACM, 2011.

[11] R. M. Sapolsky. *Why Zebras Don't Get Ulcers*. Times Books, New York, 3rd edition, 2004.

[12] C. Scollon, C.K.-Prieto, and E. Diener. Experience sampling: Promises and pitfalls, strengths and weaknesses. *J. Happiness Studies*, 4(1):5–34, 2003.

[13] S. Taylor, W. Welch, H. Kim, and D. Sherman. Cultural differences in the impact of social support on psychological and biological stress responses. *Psychological Science*, 18(9):831–837, 2007.

[14] K. Yoshiuchi, Y. Yamamoto, and A. Akabayashi. Application of ecological momentary assessment in stress-related diseases. *BioPsychoSocial Medicine*, 2(1):13, 2008.