

Singapore Management University Institutional Knowledge at Singapore Management University

Research Collection School Of Information Systems

School of Information Systems

10-2015

Mood self-assessment on smartphones

Le Minh KHUE


Eng Lieh OUH

Singapore Management University, elouh@smu.edu.sg

Stan JARZABEK

DOI: <https://doi.org/10.1145/2811780.2811921>

Follow this and additional works at: https://ink.library.smu.edu.sg/sis_research

 Part of the [Communication Technology and New Media Commons](#), and the [Software Engineering Commons](#)

Citation

KHUE, Le Minh; OUH, Eng Lieh; and JARZABEK, Stan. Mood self-assessment on smartphones. (2015). *WH 15: Proceedings of the conference on Wireless Health, Bethesda, Maryland, 14-16 October*. 1-8. Research Collection School Of Information Systems.

Available at: https://ink.library.smu.edu.sg/sis_research/4017

This Conference Proceeding Article is brought to you for free and open access by the School of Information Systems at Institutional Knowledge at Singapore Management University. It has been accepted for inclusion in Research Collection School Of Information Systems by an authorized administrator of Institutional Knowledge at Singapore Management University. For more information, please email libIR@smu.edu.sg.

Demonstration Paper: Mood Self-Assessment on Smartphones

Le Minh Khue

Department of Computer Science,
School of Computing
National University of Singapore
lmk19922000@gmail.com

Eng Lieh Ouh

Department of Computer Science,
School of Computing
National University of Singapore
issoel@nus.edu.sg

Stan Jarzabek

Faculty of Computer Science
Bialystok University of Technology,
Poland
s.jarzabek@pb.edu.p

ABSTRACT

Mood has been systematically studied by psychologists for over 100 years. As mood is a subjective feeling, any study of mood must take into account and accurately capture user's perception of an experienced feeling. In last 40 years, a number of pen-and-paper mood self-assessment scales have been proposed. Typically, a person is asked to separately rate various dimensions of the experienced feeling (e.g., pleasure and arousal) or mood items (interested, agitated, excited, etc.) on numeric scales (e.g., between 0 and 10). These partial ratings are then combined into an overall mood rating (or into its positive and negative affect). Pen-and-paper mood scales are used in basic research on mood and in clinical practice. Mobile technology makes it possible to extend mood self-assessment from lab to real life rather, collecting mood data frequently, over long time, in variety of life situations. With these motivations, we developed mobile versions of validated pen-and-paper scales for mood self-assessment to facilitate accurate in-situ mood self-assessment in real-life situations by smartphone users. The novelty of our Mobile Mood Scales (MMS) app is the use of visual effects such as color, changing brightness, animation and photos. We believe these mobile-technology-enabled aids involving user's senses can make mood self-assessment more intuitive and engaging for users than pen-and-paper mood scales that rely on linguistic terms and numerical rating. We built a customization layer that allows a doctor to generate a required mood app by selecting the mood scale required (e.g., PANAS or SPANE) as well as specific optional features such as the granularity of a rating scale (e.g., 5-point scale with radio buttons) and visual effects. In an evaluation survey, 61% of 48 participants found special features such as use of color, brightness and photos helpful in reflecting on own mood. 83% of 48 participants preferred mobile mood scales over pen-and-paper scales. We received encouraging feedback from the designers of original pen-and-paper mood scales. We envision applications of MMS in psychological studies of mood, in monitoring the efficacy of medical interventions and medication, as a component for mHealth apps where it is important to know fluctuations of patient's mood.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.

WH '15, October 14-16, 2015, Bethesda, MD, USA © 2015 ACM. ISBN 978-1-4503-3851-6/15/10...\$15.00

DOI: <http://dx.doi.org/10.1145/2811780.2811921>

Categories and Subject Descriptors

J.3 [Life and Medical Sciences]: Health; J.4 [Social and Behavioral Sciences]: Psychology

General Terms

Design, Experimentation, Evaluation, Human Factors, Measurement

Keywords

Mood scales, mood self-assessment, mobile technologies, experience sampling

1. INTRODUCTION

With fast evolution of mobile technology, smartphones are becoming a potential tool for collecting user data that can be useful in medical treatment, self-monitoring for personal improvement and many other possible applications. Among user data, mood is most elusive, difficult to capture, but at the same time important in devising intelligent human-computer interactions, affective computing [20], and user-aware mobile apps. In mobile apps for healthcare (mHealth), knowing the fluctuations of patient's mood patterns may be useful in patient monitoring, as an indicator whether a given medicine or medical intervention works for a patient or not.

As mood is a subjective feeling, any study involving mood, emotion or consciousness in general must take into account and accurately capture user's perception of an experienced feeling. In last decades, psychologists developed a number of pen-and-paper self-assessment mood scales that have been empirically validated. A typical mood scale includes from two (core affect [7][19]) to 60 (extended PANAS [24]) mood items that are believed to collectively characterize mood. These mood items are rated separately, typically on five-point scales (seven- and nine-point scales are also used in mood studies). An overall mood (or its positive and negative component) is then computed from separately rated mood items. Mood scales are used in clinical practice, surveys, basic studies on mood in psychological research, and in attempts to build mood engines to infer mood from user behavior captured by smartphones [9][14].

There are good reasons to equip smartphones with mood self-assessment capability. Pen-and-paper scales use linguistic terms and numeric ratings. Mobile technology creates interesting options to make mood self-assessment more intuitive. We may use color, changing brightness to reflect the intensity of the feeling, animation, visual effects, photos, sound or haptics (pressure or vibrations) to engage human senses that are more closely linked to mood than linguistic terms or numbers. Such

mood reporting can be more intuitive therefore possibly more accurate than with pen-and-paper mood scales – a hypothesis that should be validated in future studies. Pen-and-paper life satisfaction and mood surveys are expensive to conduct and are usually limited to small sample of participants. Similar surveys with smartphones could include large population, at low cost. If mood self-assessment via smartphones can be as accurate as pen-and-paper mood questionnaires, then smartphone-based surveys could lead to observations and findings beyond what was possible to achieve before with pen-and-paper surveys. Unlike pen-and-paper surveys, smartphones can collect in-situ mood self-reports, many times during the day, over extended period of time. This also opens the possibilities of new types of mood studies, yielding new results compared to what was possible with pen-and-paper methods.

Participants in mood studies may find entering mood frequently and for longer time tedious and boring. Smartphones can add an element of engagement via social connectivity, feedback, collaboration or competition (using gamification strategies) to counter such risks.

We believe these mobile-technology-enabled strategies involving user's senses can make mood self-assessment more intuitive and interesting for users than pen-and-paper mood scales. We hope mobile mood scales can achieve high accuracy of mood ratings, despite inconveniences and risks of having users rate their mood in real-life situations, without assistance of trained instructors.

With the above motivations and goals in mind, we developed Mobile Mood Scales (MMS) app running on Android 4.1 or higher described in this paper. MMS implements validated mood scales of PANAS [24], short PANAS [22], SPANE [4], PAM [9], SAM [2] and AffectButton [3]. While MMS can be used as a standalone app for mood tracking, our prime goal was to provide psychology researchers conducting studies on mood with a package of mood scales that can be flexibly adapted to their specific needs. Other applications of MMS include mHealth apps where it is important to know fluctuations of patient's mood, large-population low-cost mood and well-being surveys, research on mood inferring [9], and in Affective Computing [17]. MMS can be adapted to multiple application contexts. In particular, mHealth researchers and app developers can select a mood scale required in their mHealth app or experiment, as well as specific optional features such as the use of visual effects for that scale.

We received encouraging feedback on MMS from the designers of original pen-and-paper mood scales. In an evaluation survey of mobile mood scales, 61% of 48 participants found special features such as use of color, brightness and photos helpful in reflecting on own mood. 83% of 48 participants preferred mobile mood scales over pen-and-paper scales. Validating accuracy of mobile scales was not in the scope of the study reported in this paper. We plan to conduct such validation in experiments that will involve psychologists having extensive experience with conventional pen-and-paper mood scales.

In Section 2, we discuss pen-and-paper scales and other related work. We describe mobile mood scales in Section 3, and discuss their evaluation in Section 4. We comment on customization in Section 5 and Conclusions end the paper.

2. BACKGROUND AND RELATED WORK

Emotions are short-lived feelings that rapidly arise in reaction to some object or event. Unless prolonged by our mind activity (story lines), they quickly get transformed or disappear. Emotions

help us interpret the world and shape our reactions to life events. Mood is a background feeling that lasts longer than emotion – hours, days, or weeks. Most of the scales implemented into MMS can be used for self-assessment of either emotions or mood. To assess emotion the scoring question may be 'rate how you feel right now', while for the mood 'rate how you have been in last week'. In this paper, we refer to the assessed feeling using term mood.

There has been much research in recent decades on sensing emotion from physiological data collected via sensors, facial expressions, eye movement, behaviour, voice or written documents. This work focused more specifically on emotional reactions to events than on mood. Mood self-assessment methods complement rather than compete with research on automated emotion sensing. Any study of affect at some stage must consider first person's account of the affect, as there is no direct third party observer for the affect a person experiences. All visible or detectable cues have cultural, linguistic and personal bias.

2.1. Overview of mood self-assessment scales

Several mood models and mood self-assessment scales have been proposed by psychologists through the decades. Typically, to assess the mood a person rates separately a number of mood items (or dimensions) which are combined into an overall mood score. Some mood scales offer a long list of items to be rated. While this may lead to accurate mood assessment, it also poses a practical problem of time and effort required for mood rating. Yet other mood scales may involve mood items that are not easily understood without a proper training. More recent proposals try to maximize ease of use, while providing satisfactory levels of accuracy.

Circumplex model measures core affect that was found to be universally present in human emotions, and distinctly felt in our conscious experience [19][21]. Core affect decomposes mood along the pleasure and arousal (or activation) dimensions. The pleasure dimension measures how positive or negative a person feels. The arousal dimension measures how active a person is. Each mood can be understood as a linear combination of these two dimensions, or as varying degrees of both pleasure and arousal.

The PAD (Pleasure-Arousal-Dominance) model [11] adds the third dimension of Dominance to the core affect. PAD defines three nearly independent axes of the mood space.

Core affect and PAD provide a powerful framework for representing mood and have been used in many mood studies. Still, direct rating mood on three PAD scales is not easy for participants, and usually requires training and assistance of expert instructors [6].

Semantic Differential Scale (SDS) [12], the first attempt to address the above problem, consists of six pairs of bipolar adjectives related to each PAD dimension, presenting a total of eighteen ratings. Validity of SDS for mood assessment was empirically confirmed, however providing six ratings for each of the three PAD dimensions was found time-consuming, and was generating hefty dataset for statistical analysis. SDS's verbal rating system limited its applicability to native English populations.

The Self-Assessment Manikin (SAM) [2] is a visual scale for measuring PAD. Each PAD dimension is associated with a series of manikins that express the rating levels for that dimension.

Manikins change facial features, the energy level visualization, and size to enable intuitive ratings of PAD dimensions. Pictorial rather than verbal scales make SAM accessible to non-native English speakers. A computer-based version of SAM is also available. Our mobile mood scales to much extent were inspired by SAM.

More recently, there have been attempts to make use of computers to measure mood instead of traditional pen-and-paper methods. One prominent work is AffectButton [3]. AffectButton is a PC tool for emotion self-reporting. It relies on PAD model to measure mood. The user moves the mouse over the graphical image of a human face that forms AffectButton's user interface. The face changes the expression and the task of the user is to identify an expression that matches her mood. Each face expression shown by AffectButton corresponds to a unique combination of PAD values. AffectButton is simple and intuitive to use. This is very important in studies when participants are prompted to submit mood self-reports in daily situations. AffectButton has not been validated as extensively as other mood scales, but we can safely say that using AffectButton we can get more accurate mood assessments than just by asking users whether they feel happy or sad, because the meaning of these words may be different from person to person. In contrast, a facial expression may be clearer to them. Lastly, using AffectButton is more fun than using scales involving linguistic and numeric terms.

Cited over 7000 times, the Positive and Negative Affect Schedule (PANAS) model [24] is based on the idea that people can feel good and bad at the same time [8]. Hence, PANAS uses two scales to measure Positive Affect (PA) and Negative Affect (NA), respectively. Each scale consists of 10 mood items such as irritated, inspired, etc. For each of the items, user can rate the extent to which she experiences the corresponding feeling (lowest is 1: 'very slightly or not at all' and highest is 5: 'extremely'). The PA and NA mood characteristics can be computed from the individual ratings by performing a principle factor analysis [24].

According to Thompson [22], the original version of PANAS is quite long for studies in which participants are asked to report mood many times per day. Furthermore, the full 20-item PANAS contains some words that are ambiguous. A 10-item PANAS short form has been introduced to address the above shortcomings. Five of the 10 items measure PA and the other five items measure NA.

Similar to PANAS, the Scale of Positive and Negative Experience (SPANE) [5] is designed to assess subjective feelings. SPANE comprises 12 items, six of which assess positive feelings and the other six items assess negative feelings. Unlike PANAS, SPANE includes general feelings, such as 'positive', 'negative' and 'pleasant', which allows it to reflect a full range of mood and feelings without creating a list of many items to fully reflect the diversity of positive and negative feelings [5]. Therefore, ratings of individual items in SPANE can be combined to compute the overall mood rating. Lastly, SPANE does not contain items such as 'active' and 'strong' that do not refer to feelings and items such as 'jittery' that have been found to be difficult to interpret for non-native English speakers. Each SPANE item is scored on a scale ranging from 1 to 5, where 1 represents 'not at all' and 5 represents 'extremely'.

Addressing the issue of lengthy questionnaire-based mood scales, the Photographic Affect Meter (PAM) is designed to run on modern mobile phones and can support in situ affect measuring [18]. Users of PAM are shown a grid containing photos of varying subject matter, each in some way depicting a mood state. Each

photo is associated with a mood name. The arrangement of photos in PAM is relied on the Circumplex model. Users are prompted to select a photo that best describes how they feel right now. Users can reload the set of photos in case they are unable to find the photo that matches their feeling.

2.2. In situ self-reporting

In situ mood assessment is more difficult than mood assessment in lab-based experiments. Diary techniques [20] have been commonly used for in situ mood assessment. Participants record their daily activities or mood on a preprinted log form which is then returned to researchers. A number of problems have been identified with diary approach. Firstly, participants tend to fill log form after the event which adds cognitive bias to the reported data. Secondly, it is difficult for researcher to gain access to the study data before the participants submit these forms to them. Finally, letting participants keep these forms for a long time may affect the integrity of data as participants may get tempted to modify the log forms of the previous days.

Along the line of diary techniques, Palen and Saltzman [16] introduced the voice-mail diary method, where participants use mobile or landline phones to make reports to a dedicated voice-mail line instead of recording events on paper. This method simplifies reporting and solves the issue with data integrity faced by the traditional pen and paper diary. However, one limitation is that the receiving voice-mail line cannot handle simultaneous incoming calls. Therefore, if one participant is calling the line, others need to wait. Another disadvantage of this method lies with the fact that the data collected is in the audio form, not in a format that can be easily analyzed by common analysis tools.

Modern technology allows researchers to use smartphones as a tool to collect in situ self-reports. Healey [6] and Morris et al. [13] developed journaling system and Mood Map app to collect mood information from participants, based on variations of PAD and Circumplex models.

3. MOBILE MOOD SCALES (MMS) APP

MMS app is a package of multiple mood scales on Android (4.1 or higher) smartphones. Users can customize MMS by selecting required mood scales and their features.

3.1. Tour through MMS app

The reader may choose to view the two demo videos of mobile mood scales before reading this section:

<http://www.youtube.com/watch?v=OSTcwGizkgo&feature=youtu.be>

https://www.dropbox.com/s/a8j29rxfx1iweyl/Padsam_video_demo.wmv

MMS app supports PANAS (20-item version), PANAS short form (10-item version), SPANE, AffectButton and PAM. PANAS, PANAS short form and SPANE mood scales are implemented in versions that closely reflecting their original pen-and-paper counterparts, and in enhanced versions that employ pictures and visual effects. Customization options allow the user to choose a required form of the scale. The MMS screen snapshots of scales below demonstrate various options available to the user. For example, we show mobile SPANE in the original form, with numeric scales for rating 12 mood items, and PANAS enhanced with pictures.

Users of SPANE (Figure 1) rate each of the 12 mood items on the scale from 1 to 5 using radio buttons. Pressing the icon in the bottom left corner of the screen saves the mood recording to the smartphone's memory.

While reporting mood, the user may comment on any special circumstances at the time when the mood rating was taken, by clicking the icon in the middle bottom of the screen. For example, if the user indicates that she is in a happy mood, she can note that she just received a gift from friends. The icon on the right bottom of the screen connects the user to friends or the doctor.

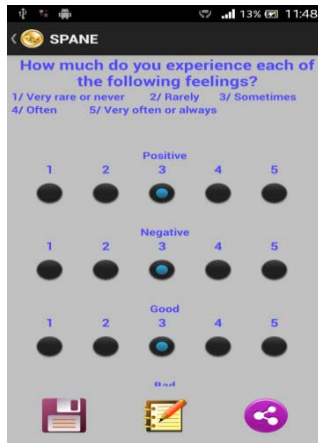


Figure 1. SPANE in MMS

In PANAS (Figure 2), we show original scales enhanced with photos. Each of the 20 mood items is decorated with a photo that reflects the meaning of that item. Photos help user in case there is an ambiguity in the meaning of the items with the brightness of the photos changes according to user's rating of each item, where 1 means lowest brightness and 5 means highest brightness.

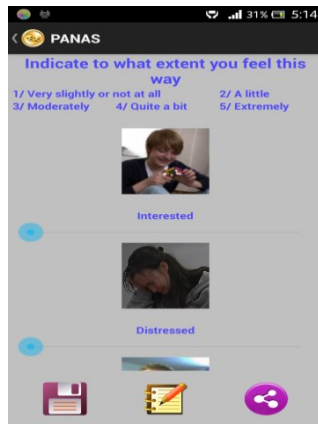


Figure 2. PANAS in MMS

There is much empirical evidence on picture-evoked affects that suggests that pictures and photos can have universal emotional legibility [7]. We first considered the International Affective Picture System (IAPS), a collection of photos that makes up a validated instrument for evoking a variety of emotional responses [7]. However, as IAPS's purpose is to elicit rather than to describe mood, we thought that by using IAPS photos in MMS we could be actually affecting the mood ratings. Eventually, we chose to use the online photo-sharing service Flickr (<https://www.flickr.com/>) under non-commercial license. We looked for the most popular photos with the same name tag as items in PANAS, PANAS short form and SPANE. We then conducted a focus group to select the best photo for each item.

Another feature variation that we see in mobile PANAS of Figure 2 is a rating scale with a slider rather than a radio button as in the

SPANE scale. For any mood scale, the user may choose either radio buttons or a slider.

Figure 3 shows the MMS user interface for the AffectButton. As the user moves the finger around the face icon, the expression changes accordingly showing a range of emotions. As before, icons at the bottom allow the user to save the mood recording, make a note and share mood with friends or doctors.

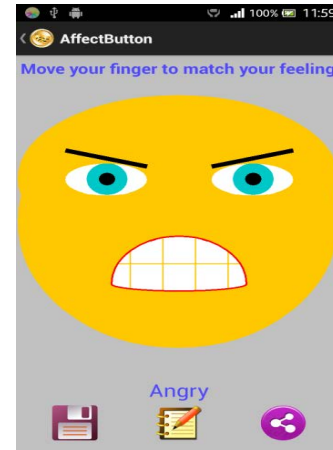


Figure 3. AffectButton in MMS

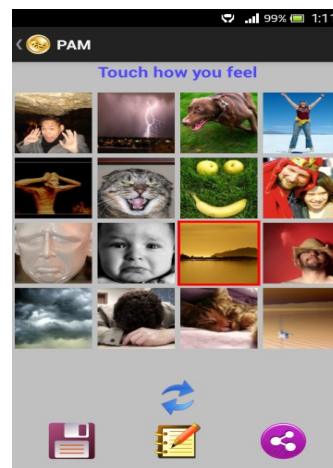


Figure 4. PAM in MMS

PAM scale is shown in Figure 4 with one of the sets of photos for mood selection. The set of photos can be changed in case the user cannot comfortably identify her mood in a given set.

In SAM, PAD mood dimensions are decorated with a series of human-like manikins that express the rating levels for that dimension. Visualization of pleasure, arousal and dominance mood aspects makes mood assessment more intuitive as compared to verbal explanations. Our mobile version of SAM provides the following customizable features that intent further enhance intuition, relatedness and engagement of scale users:

- Choice between original SAM manikin and a more human-like figure such as shown in Figure 5,
- Choice of male or female human-like figure and male, female or neutral for SAM manikin,
- Raising arms and changing brightness as the intensity of the feeling increases on the scale,
- Use of color.

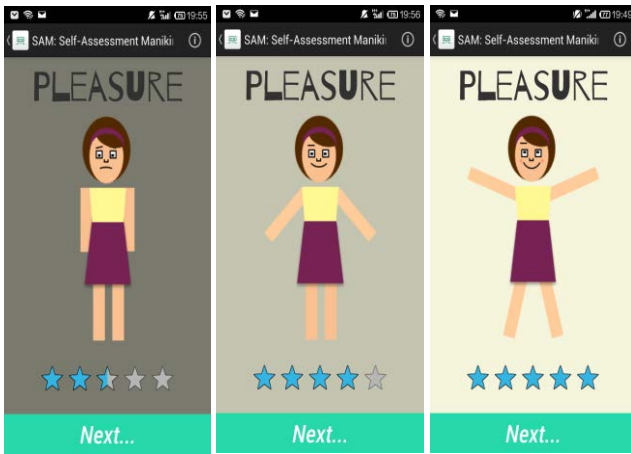


Figure 5. Mobile version of SAM

Three female human-like figures in Figure 5 show three ratings of pleasure dimension with intensity increasing from the left to the right.

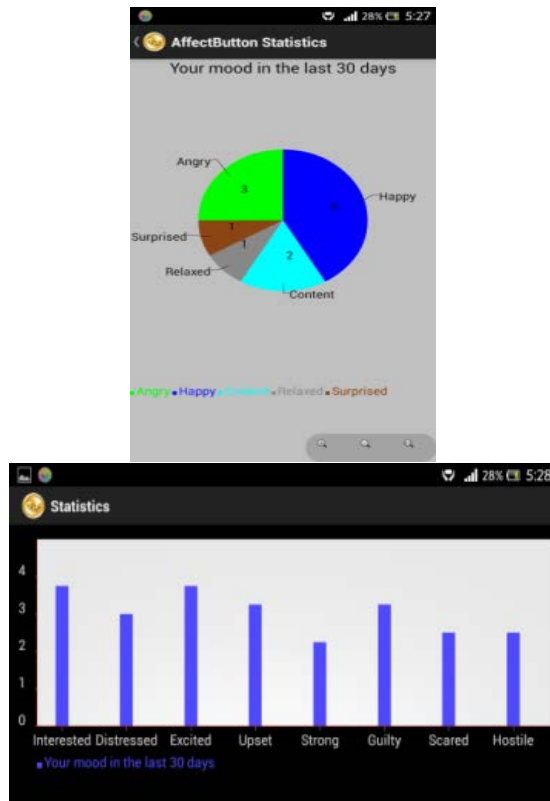


Figure 6. Graphs of mood history in MMS

Users can view their mood history through pie charts or bar charts as shown in Figure 6. These charts can be zoomed in or out flexibly based on user's preference. Since MMS is an app primarily for research purpose, participants in experiments may be required to send their mood data to researchers regularly. Export feature allows the user to store a detailed report of their mood history in PDF or CSV formats, and email to a researcher as an attachment.

Figure 7 shows how the user can share her mood with friends via Facebook, Google+, email, SMS, etc.

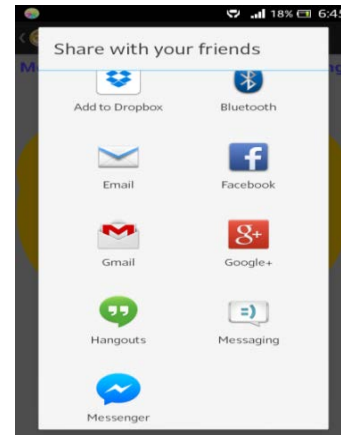


Figure 7. Sharing mood feature in MMS

4. EVALUATION

The novelty of our Mobile Mood Scales (MMS) app is the use of photos, visual effects, multiple forms of rating scales, social connectivity and graphical views to feedback historical mood data to the user. We designed MMS to engage the user in mood reporting, making it possible to collect mood data frequently, for extended time. We conducted two studies to evaluate mobile scales and their various novel features that distinguish them from the original validated pen-and-paper scales.

4.1. Initial ad hoc survey

The aim of this study was to collect quick early feedback about the usability of mobile mood scales at the final stage of their implementation. We asked 40 students to freely play with our app and provide their feedback. Survey questions were related to general utility and usability of the app; usefulness of various features such as graph visualization of mood history or sharing mood with friends; and customizable settings of the app. We also asked participants to compare their experience using various scales and assess their engagement when using mood scale app, an important quality for apps in general [1], and particularly important in mood self-assessment experiments that require collection of mood data frequently and over extended time.

We employed the System Usability Scale (SUS) [4]. SUS consists of 10 statements with five response options for each statement (1 means "Strongly disagree" and 5 means "Strongly agree"). Our survey included seven items adapted from the User Engagement Scale [15]. Participants rated engagement on a five point scale as in the first part of the survey.

The reader can view our survey questionnaire at:

<https://docs.google.com/forms/d/1Xxp4xAddSVLlbkgdu4IIMIREmVywWQy8HDiOYvJ3UVE/viewform>

The results: App features that have to do with intuitive mood self-assessment have received ratings above 3 on a 5-point scale. This result confirmed our expectation that using smartphone technology mood reporting can be interesting and engaging. As expected, a higher proportion (45%) of participants preferred AffectButton as the method to record the mood. However, in terms of accuracy of mood description, only 30% of participants thought that it was the best method to use. 73% of participants agreed that photos accompanying mood items in the scales helped them better relate to their mood. However, only 38% of participants thought that the changing brightness of photos was useful. Hence, in the future development of the app, we need to

further study this aspect of mood reporting and consider alternatives. A potential alternative is to use phone vibration as an indicator of intensity associated with a given mood item: Weak vibration corresponding to low intensity and increasingly strong vibration corresponding higher intensity. We expected that participants would prefer a slider over radio buttons for mood rating scales. However, participants did not think the form of the scale was important. We received an interesting feedback suggesting that the color scheme of the themes should be carefully selected since it may affect the current mood of user at the time of self-reporting and cause her to lose her objective judgment.

We also learned a bitter lesson regarding the importance to properly motivate and explain the intended usage of the mobile mood scale app to participants. Participants wrongly assumed that in order to assess their mood they need use all the mood scales at the same time. In reality, only one mood scale is used at any given time. The post-survey discussion with participants revealed this misunderstanding and we had to discard part of the survey related to utility and engagement that was affected by the incorrect assumption. We realized that we failed provide participants with detailed enough protocol how to use our app during evaluation.

4.2. A systematic evaluation

Encouraged by generally positive feedback, we thought our mood scale app deserved a more thorough evaluation, free of flaws encountered in the initial survey. We designed a new evaluation study for which we recruited 48 participants, comprising of graduates from Master and Graduate Diploma of System Analysis programmes at Institute of System Sciences, students from other faculties at NUS, and one faculty member from Kathmandu University, Nepal. The survey design follows the guidelines of ESOMAR/GRBN Guideline for online sample quality [25]. We validated the participant’s identity through an initial survey briefing (face-to-face) with no personal identifiable data being shared without the consent of the participant. The survey involves participants only in the age range above 18 with no involvement of children or young people and the ratio of males over females is estimated at 60% to 40%. The survey questions and inputs are automatically collected using Google Docs in a spreadsheet format.

The study protocol: Our new study spread over 11 days. Before the study, we briefed the participants on the study protocol and introduced them to mobile health and also the current assessment method using paper mood scales. During first 10 days, participants used a specific configuration of one mood scale per day. They entered mood twice: the first time primarily concentrating on observing own mood, and the second time giving prime attention to app’s features and user interface. We asked participants to fill the survey just after they had entered the mood for the second time on a given day. On the final day, the participants used pen-and-paper mood scales and we asked them to compare their impressions of using both modes of mood self-assessment. Each survey (both using mood scale and filling survey questions) only take less than 10 minutes to minimize break-off or disengage of the participant, potentially jeopardize data quality. We did not conduct the pen-and-paper mood scale over another 10 days due to concern of survey fatigue. Our approach is to introduce the pen-and-paper mood scales to the participants during survey briefing and again on day 11 after all the mobile mood scales to make their concluding remarks for pen-and-paper versus mobile mood scales. As in the initial study, we based survey questions on SUS [4] and User Engagement Scale [15], with answers rated on 5-point scale from 1 - “Strongly

disagree” to 5 - “Strongly agree”. Because we judged that validating mobile scales against the original pen-and-paper mood scales need be done in a separate study, we did not collect the actual mood data to avoid privacy concerns.

The website and survey: We provided solid motivation and background for participants to understand the reasons behind the survey in the pre-study meeting and on the website <https://sites.google.com/site/surveyofmsaandpadsam/>.

The results: In our survey, the participation rate (based on definition of ESOMAR/GRBN [25]) stands at 100% with completion rate at 98% (1 only complete 60% due to device incompatibility). The number of duplicate responses of each participant out of a total of 517 usable responses is 6 (98.8%). There were 5 questions related to the ease of use of various forms of mood scales, 2 questions related to the level of engagement, and 2 questions to rate the ease at which participants found a given scale helpful to relate to their mood. The results are shown in Figure 8.

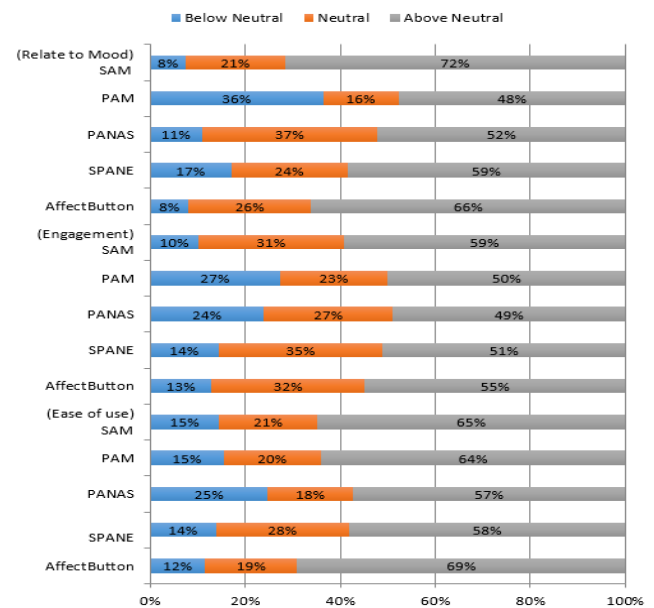


Figure 8 Usability and Engagement of Mobile Mood Scales

The survey included 7 questions to let participants evaluate how special features such as the use of color, brightness, photos or animation helped them in mood rating. The summary of answers is shown in Figure 9.

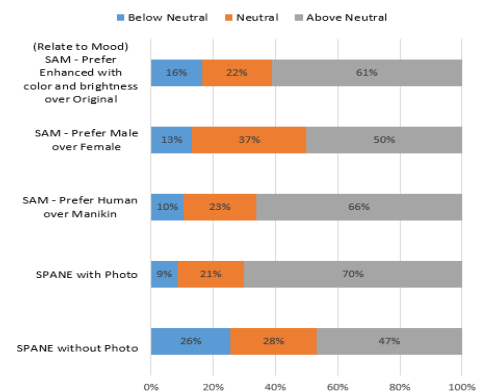


Figure 9 User preference for special features

We also evaluated features such social connectivity or graphs showing mood data participants entered during the study (Figure 10). There were 6 questions in this category.

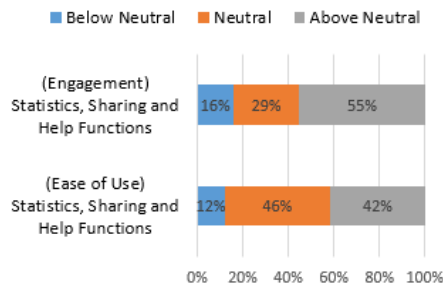


Figure 10 Engagement and ease of use

In the last part of the questionnaire, we asked participants to use original pen-and-paper mood scales and compare their experience with mobile mood scales. We also asked the participants to rate their preferred mobile mood scale. There were 5 questions in this category and the results are shown in Figure 11 and Figure 12.

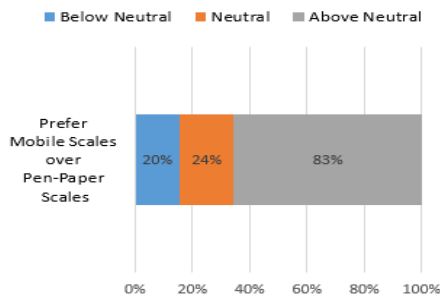


Figure 11 Mobile Scales versus Pen-Paper Scales

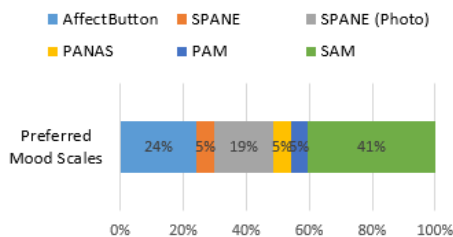


Figure 12 Preferred mobile mood scales

Observations: Participants favored SAM over other scales for all the evaluation criteria. Furthermore, they favored human-like figure over abstract icon-manikin used in original SAM. Most of the participants favored male figure independently of their own gender. Participants found photos helpful in relating them to their mood, as shown in comparison of SPANE with and without photos (Figure 9).

SAM and AffectButton received highest ratings in terms of engagement and ease of use. Viewing past mood data, social connectivity and help functions were found engaging, however there is room to improve ease of use of apps.

83% of participants preferred mobile scales to pen-and-paper scales with SAM as the overall preferred mobile mood scale as shown in Figure 11 and Figure 12. 45% of participants indicated gender of icons in mood scales to be of considerable importance in helping them relate to their mood. In most comments we read that participants preferred icons of their own gender.

As a summary of the survey, we extracted three quotes from the participants: (1) “Although I felt easier to just do a short mood survey once on pen-and-paper, I would rather do it on a mobile if on a daily basis for the interactivity.”, (2) “It is quite difficult to rate the mood on pen-and-paper only based on terms or static pictures.”, (3) “Paper work is boring unlike using the app which is more interesting and engaging with its colorful features and animation.”.

5. CUSTOMIZABLE PACKAGE OF MOBILE MOOD SCALES

Initially, we implemented mobile mood scales in two apps with the user being able to select a required scale from the menu. Our goal was to evaluate usefulness and appropriateness of various features of mobile scales. However, only one or two mobile scales would be used in the context of real applications such as mood data collection in well-being surveys or in psychological research on mood. The latter is the main intended application of mobile mood scales. Other than selecting a required mood scale, the mood study designer may want to select a required combination of features such as the type of a rating scale (slider, radio button or stars), the number of rating points (ranging between 5 and 100), the use of brightness, pictures, etc. For the convenience of doing customization, we implemented a web application that provides intuitive interface for the user to specify required moods scales and their features. A custom mood scale app is then automatically produced and can be downloaded to smartphones.

We used ART (Adaptive Reuse Technique art.comp.nus.edu.sg) for flexible customization of mood scale apps. ART allowed us to partition mood scales and their features into distinct modules parameterized for ease of adaptation. ART engine performs synthesis of a required mood scale app from specifications of mood scales required and their features. In the process of this synthesis, app modules are composed after suitable customizations to form an executable app that can be uploaded to the users’ smartphones.

6. CONCLUSION AND FUTURE WORK

We have presented mobile versions of validated pen-and-paper mood assessment scales, including PANAS, PANAS short form, SPANE, AffectButton and PAM, implemented as an Android app called Mobile Mood Scales (MMS). Our intention was to experiment with the idea of using smartphones for in situ mood self-assessment. Mood rating in MMS is aided by visual effects such as changing brightness, animation, photos, social connectivity and graphical views to feedback of historical mood data to the user. Our motivation was that by involving user’s senses mood self-reporting could be more intuitive and interesting. In that way, we hoped to achieve high accuracy of mood ratings, despite inconveniences and risks of having users rate their mood in real-life situations, without assistance of trained instructors. We tried to apply mobile strategies to engage the user in mood self-reporting, so that mood data could be collected frequently, for extended period of time as it is required in many mood studies. For PANAS, PANAS short form, SPANE, and SAM to our best knowledge ours is the first attempt to bring pen-and-paper scales to mobile phones.

MMS provides a platform for psychological studies of mood and emotion in real-life settings and for comparative studies of various mood scales. Other applications of MMS include mood tracking, as a mood self-assessment component for mHealth apps where it is important to know fluctuations of patient’s mood or to track the impact of a drug on person’s mood, in large-population low-cost

mood and well-being surveys, and in research on inferring mood from smartphone user behavior. We designed MMS as a customizable package of mood scales to ensure ease of adaptability to these multiple application contexts. mHealth researchers and app developers can customize MMS by selecting a mood scale required for a given mHealth app or mood study, as well as specific optional features such as the use of visual effects for that scale.

In an evaluation survey, we focused on usability of mobile mood scales, and their ability to help users relate to the experienced mood. We also checked user perception of mobile versus pen-and-paper mood scale. Among the results, 61% of 48 participants found special features such as use of color, brightness and photos helpful in reflecting on own mood. 83% of 48 participants preferred mobile mood scales over pen-and-paper scales. Mobile SAM received highest ratings in respect to all our evaluation criteria.

The work reported in this paper is just the first step in exploring mobile mood self-assessment methods. The mood rating accuracy of mobile mood scales must be validated in controlled experiments by comparing them with already validated pen-and-paper mood scales. Each group of customizable features of mobile scales should be also validated in separation from others to see their impact on accuracy of mood ratings. An important question is to what extent special features such as photos or changing brightness help people reflect on their mood and to what extent they also affect the mood rating. We will team up with psychologists to conduct these evaluations.

7. ACKNOWLEDGEMENT

This work has been supported by Microsoft Asia Research Grant R-252-000-521-592. We are grateful to William Tov and David Watson for giving us insightful feedback on early version of mobile mood scales. We would like to thank NUS student Eugene Sim for the implementation of mobile SAM scale and Howard Russon for assisting to reach out to students to participate in the detailed study.

8. REFERENCES

- [1] Bannon, L. 2005. A human-centered perspective on interaction design. *Future Interaction Design*, Springer (2005), 9-30.
- [2] Bradley, M. and Peter J. Lang, P. 1994. Measuring emotion: The Self-Assessment Manikin and The Semantic Differential. *J. Behav. Ther. & Exp. Psychiat.* 25, 1 (Mar. 1994), 49-59.
- [3] Broekens, J. and Brinkman, W.P. 2009. AffectButton: Towards a Standard for Dynamic Affective User Feedback. *Affective Computing and Intelligent Interaction (ACII)* (2009).
- [4] Brooke, J. 1996. *SUS: a "quick and dirty" usability scale* (1996).
- [5] Diener, E., Wirtz, D., Tov, W., Chu Kim-Prieto, Dong-won Choi, Shigehiro Oishi, S. and Biswas-Diener, R. 2009. New Well-being Measures: Short Scales to Assess Flourishing and Positive and Negative Feelings. *Social Indicators Research*. 97, 2, 143-156.
- [6] Healey, J. 2011. *Recording Affect in the field: Towards Methods and Metrics for Improving Ground Truth labels*. *Affective Computing and Intelligent Interaction*. 6974 (2011), 107-116.
- [7] Lang, P. 1995. The Emotion Probe: Studies of Motivation and Attention. *American Psychologist*. 50, 5 (1995), 372-85.
- [8] Larsen, J.T., McGraw A.P. and Cacioppo, J.T. 1988. Can people feel happy and sad at the same time? *Journal of Personality and Social Psychology*.
- [9] LiKamWa, R., Liu, Y., Lane, N.D. and Zhong, L. 2013. MoodScope: Building a Mood Sensor from Smartphone Usage Patterns. *MobiSys*, (Jun 2013), 389-402.
- [10] Measuring Usability: <http://www.measuringusability.com/>
- [11] Mehrabian, A. 1996. Pleasure-Arousal-Dominance: A general framework for describing and measuring individual differences in temperament. *Current Psychology Developmental, Learning, Personality*. 14, 4 (Winter 1996), 261-292.
- [12] Mehrabian, A. and Russell, A. J. 1974. *An Approach to Environmental Psychology*. Cambridge: M.I.T. Press (1974).
- [13] Morris, M.M., Kathawala, Q., Leen, T.K., Gorenstein, E.E., Guilak, F., Labhard, M., Deleeuw, W. 2010. Mobile Therapy: Case study evaluations of a Cell phone application for Emotion Self-Awareness. *Journal of Medical Internet Research* (April 2010).
- [14] Novak, G., Carlson, D. and Jarzabek, S. 2014. An Adaptable and Extensible Mobile Sensing Framework for Patient Monitoring. *IEEE 9th Int. Conf. on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP)*, Singapore (Apr. 2014).
- [15] O'Brien, H.L., Toms, E.G. 2010. *The Development and Evaluation of a Survey to Measure User Engagement*. *Journal of the American Society for Information Science and Technology*. 61, 1 (2010), 50-69.
- [16] Palen, L. Marilyn Salzman, M. 2002. Voice-Mail diary studies for naturalistic data capture under mobile conditions. *ACM Conference on Computer supported cooperative work* (2002), 87-95.
- [17] Picard, R.W. 1997. *Affective Computing*, MIT Press, Cambridge (1997).
- [18] Pollak, J.P., Adams, P. and Gay, G. 2011. PAM: A photographic affect meter for frequent, in situ measurement of affect. *CHI* (2011), 725-734.
- [19] Posner, J., Russell, J.A. and Peterson, B.S. 2005. The Circumplex model of affect: An integrative approach to affective neuroscience, cognitive development and psychopathology. *Development and Psychopathology* (2005), 715-734.
- [20] Rieman, J. 1993. The diary study: A workplace-oriented research tool to guide laboratory efforts. *ACM Conference on Human Factors in Computing Systems* (1993), 321-326.
- [21] Russell, J. 2009. Emotion, core affect, and psychological construction. *Cognition and Emotion*. 23, 7 (2009), 1259-1283.
- [22] Thompson, E.R. 2007. Development and Validation of an Internationally Reliable Short-Form of the Positive and Negative Affect Schedule (PANAS). *Journal of Cross-Cultural Psychology*. 38, 2 (Mar. 2007), 227-242.
- [23] Adaptive Reuse Technique: <http://art.comp.nus.edu.sg/>
- [24] Watson, D., Clark, L.A. and Tellegen, A. 1988. Development and Validation of Brief Measures of Positive and Negative Affect: The PANAS Scale. *Journal of Personality and Social Psychology*
- [25] ESOMAR World Research. ESOMAR/GRBN Guideline on Online Sample Quality <https://www.esomar.org/knowledge-and-standards/codes-and-guidelines/guideline-on-online-sample-quality.php>