Planning the Future in a Longer Perspective: Effects of a One-week Forecast of Mental Health

NAOKI TATEYAMA, RYOTA YOKOMURA, YUKI BAN, SHIN'ICHI WARISAWA, and RUI FUKUI, The University of Tokyo, Japan

A long-term perspective toward the future enables a more comprehensive approach to decision-making, considering a variety of potential scenarios. The forecasting of mental health was anticipated to promote proactive planning, however, it faces challenges such as a short forecasting period and a lack of intuitive understanding of the relationship between actions and the forecast. This study presents a novel mental health indicator that incorporates a long-term perspective by considering past actions. A four-week experiment was conducted with 105 participants to evaluate the effects of a one-week forecast. Qualitative analysis reveals the effects of the one-week forecast on behavioral planning, emotional states, and reasons for disregarding the forecasts. Findings indicate that conventional mood indicators prompt participants to prioritize pre-existing schedules and perceive the forecast as infeasible, whereas the proposed indicator enhances the ability to plan work schedules in advance. Our results offer valuable insights into the presentation of forecasts for effectively managing mental health, considering the time constraints of everyday life.

CCS Concepts: • **Human-centered computing → HCI theory, concepts and models**.

Additional Key Words and Phrases: forecast of future mental health, personal informatics, scheduling, calendar events, mental health

ACM Reference Format:
Naoki Tateyama, Ryota Yokomura, Yuki Ban, Shin'ichi Warisawa, and Rui Fukui. 2024. Planning the Future in a Longer Perspective: Effects of a One-week Forecast of Mental Health. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 8, 1, Article 18 (March 2024), 20 pages. https://doi.org/10.1145/3643538

1 INTRODUCTION

As a social issue [18], mental health problems have gained recognition and self-care technology based on personal informatics (PI) [14, 30] has gathered widespread attention. Advances in computer science have made data-driven forecasting of human internal states possible [19, 31, 45]. Such technology can help individuals envision the future and facilitate coping behaviors [5, 27, 28]. Forecasting is also expected to reduce future uncertainties and motivate the use of PI systems [2]. The longer the period forecasted, the easier it is to identify sources of stress and anxiety and address them proactively.

The presentation of forecasted mental health indicators has been a focus in previous studies [19, 22]. These studies have looked at how the forecast impacts behavior and cognition. However, they have limited their forecast to a maximum of two days. Managing mental health effectively requires balancing various life aspects, including work [7]. Key time management skills, such as setting goals and prioritizing tasks, help in handling multiple responsibilities [12]. One important aspect of time management is the ability to plan for a week or more [8, 34].

Authors' address: Naoki Tateyama, tateyamanaoki@lelab.t.u-tokyo.ac.jp; Ryota Yokomura, ryokomura@s.h.k.u-tokyo.ac.jp; Yuki Ban, ban@edu.k.u-tokyo.ac.jp; Shin'ichi Warisawa, warisawa@edu.k.u-tokyo.ac.jp; Rui Fukui, fukui@ra-laboratory.com, The University of Tokyo, Kashiwanoha 5-1-5 Environmental building 301, Kashiwa, Chiba, Japan, 277-8563.

2474-9567/2024/3-ART18 $15.00
https://doi.org/10.1145/3643538

Planning beyond a two-day scope is crucial; thus, a one-week planning horizon is adopted in this study to accommodate both workdays and weekends. To realize a balanced plan for mental health and daily activities, it is necessary to clarify the impact of a forecast extending beyond two days.

In addition to the length of the forecast, it is also important to consider how the forecast is understood and affects the decision on behaviors. Mental health indicators are easily influenced by daily events, and humans have a limited capacity to forecast their long-term future states. It is difficult to intuitively understand the causal relationship between past actions and mental health. If the forecast can be linked to past actions, users can understand how their actions affect their mental health and plan for the future.

The objective of this study is to understand the effect of a longer-term forecast of mental health indicators and to clarify whether the reflection of past actions in the forecast affects behaviors. To address this objective, we formulated the following two-step research questions.

- **RQ1. How is a one-week forecast of mental health indicators recognized and utilized?**
- **RQ2. If a forecast of mental health indicators can reflect past actions, how does the forecast affect behaviors?**

We developed a web application capable of linking to a daily-use scheduler and conducted experiments to present the user-specific forecast of mental health indicators. Participants used the web application in their real lives for approximately four weeks. A qualitative analysis examining the effects of the forecast was conducted.

The specific contributions are as follows:

- The implementation of a one-week mental health forecasting model based on calendar events, which is a longer period than previous studies was realized. The model is incorporated into a web application that enables users to plan their future actions based on the forecast.
- A new mental health indicator called "mental debt" is proposed, which takes into account past situations. This indicator reflects the situation of the previous week in a way that conventional mood indicators cannot.
- A four-week experiment involving 105 participants was conducted to introduce the web application into a real-life setting. The results of a qualitative analysis revealed the effects of the one-week forecast of mental health indicators and factors that prevented the usage of the forecast.

2 RELATED WORK

2.1 Forecast of mental health

In the PI focusing on mental health, technologies for measuring mental states were developed and used to alter behavior. Previous studies on mood prediction have used physiological signals such as the electrocardiogram, electroencephalogram, and pulse waves [9, 33, 44, 53] smartphone operation logs [43], activities obtained by wearable sensors [1, 29, 42], and text data [3], etc. These technologies are also applied to forecasting future states that have not yet been experienced. Taylor et al. used smartphone data and wearable sensors to forecast the next day's mood [49]. Their neural network model used inputs including skin potentials, smartphone calls, messages, and screen on/off states. Chikersal et al. used smartphone data and wearable sensors to forecast the depression scale three months ahead [10]. While these studies provide valuable insights by utilizing historical data, they inherently focus on extrapolating past trends into the future. This methodology may face challenges in long-term forecasting because it does not take into account unforeseen factors or changes that could shape future outcomes.

To enhance forecast accuracy, some studies have focused on future factors of mental health changes [45]. This approach is believed to be more user-centric compared to models relying solely on past moods. Li et al. leveraged the relationship between stress and words in social networks to develop a model that predicts stress using microblog posts related to future events [31, 32]. However, this approach may face challenges as users may be hesitant to share privacy-sensitive information [37]. Tateyama et al. utilized calendar events, which reflect user-specific situations, as a means of obtaining information without compromising privacy [48]. These events can also be conveniently imported from scheduling applications.

By using future events as inputs, it may be possible to forecast mental health indicators over longer periods with higher accuracy.

2.2 Forecasted lifelog for proactive actions in the real-life setting

Previous studies have investigated how humans accept predictions of internal states by PI systems [20]. The comprehensibility [22, 24] and credibility [46, 51] of the prediction are important factors for encouraging users to use the prediction. These studies have mainly focused on the interpretation of the prediction itself, however, few studies have investigated how users imagine their future states based on the prediction and how they change their behavior accordingly.

In contrast, forecasts of personal lifelog have been actively utilized in fields other than mental health. To encourage users to take action toward achieving a desired state, it is important to link their actions to future states. In the field of task management, remaining time for future events [47] and the feasibility of the user's schedule [35] were forecasted. They aimed to enhance the user's awareness of the importance of proactive steps toward achieving their objective. Forecasting is also used for physical health improvement. Rho et al. proposed a system informing user's forecasted weight based on information such as activity level and sleep [40]. The forecasted weight causes users to anticipate their future body shape and creates an awareness of goals. Miyake et al. proposed a system that forecasts the cumulative number of steps taken at specific time intervals of the day [36]. The forecasted number of steps assists in improving user satisfaction and motivation. In the aforementioned studies, the forecasted lifelog is perceived to reflect previous actions toward a goal.

Considering the constraints of daily life is also important for designing systems that utilize forecasts. Bilbily et al. integrated a scheduler and map, displaying the forecasted duration for moving [6]. Users can be conscious of the time-space constraints during scheduling. Lee et al. proposed a system to improve sleep schedules considering daily events [26]. It provides individually actionable sleep schedules. Considering constraints when planning their actions enables users to create more feasible and realistic plans.

To introduce long-term forecasts of mental health indicators, we designed our PI system focusing on linking actions to future states and considering constraints.

3 METHOD

This section describes the proposed PI system in detail. This system used daily event information and mood. The first two subsections describe how the event information and mental health indicators were collected and how the forecast was implemented. The latter two subsections describe the design of a new indicator to link the forecast to the user's action and the interface for real-life use.

Given the heterogeneity in background variables such as occupation and age, it is difficult to understand the effects of the forecast. This study describes a design that assumes a homogeneous group of participants in the experiment, consisting only of undergraduate and graduate students.

3.1 Lifelog collection interface

Daily event and mood information was collected using the day reconstruction method (DRM) [21]. The DRM enables participants to recall the daily events in chronological order and to reconstruct the day, thereby reducing recall bias [50]. Figure 1 (A) shows the interface for comprehensively recording events that occurred during the day. It collaborates with the user's Google Calendar via Google Calendar API, mitigating the risk of missed events. Participants can also record events that have not been registered on Google Calendar. Figure 1 (B) shows a pop-up interface for recording information for each event. Here, participants record the name and the category

Fig. 1. Interfaces for collecting participants' daily events and moods. Text on the application was displayed in the local language. (A) Record of daily events. (B) Record of event details. (C) Record of daily mood. (D) Record of total uplifting and depressive effects of events.

Table 1. Labels that can be assigned to events.

Category	**Task**: "work", "research", "part-time job", "job hunting", "trivial duties", "lesson", "study", "moving" **Activity**: "going out", "hobbies and entertainment", "exercise", "travel", "dating", "chatting" **Life**: "meals", "sleep", "rest" **Other**: "other"
Person	alone, friend, senior person, junior person, colleague, family, partner, stranger, collaborative research company, company for job-hunting, other community, other

of the event, with whom they have performed the event and the effect on their mood throughout the event. The category and person can be selected from labels listed in Table 1. Participants can choose more than a single label.

The effect of the event on the mood at the time of the event was entered through a bar graph-like interface located at the bottom of Figure 1 (B). The uplifting effect could be specified by clicking or tapping the positive part of the graph, whereas the depressive effect could be specified from the negative part. Participants specified a value for the effect ranging from 0 to 100.

Figure 1 (C) shows the interface for recording daily mood, with the slider bar ranging from "very negative" on the left to "very positive" on the right and "neither" in the center. The values recorded by the previous day's responses can also be referred to as mood criteria.

Participants reported the total effects of the day's events in Figure 1 (D). They focus on overall appraisal (cognitive evaluation) of the events which is also important for mental health evaluation [25]. Only the total uplifting and depressive effects were measured for overall impact in this study.

To provide reference values, physical vitality was also requested, but not utilized in this study.

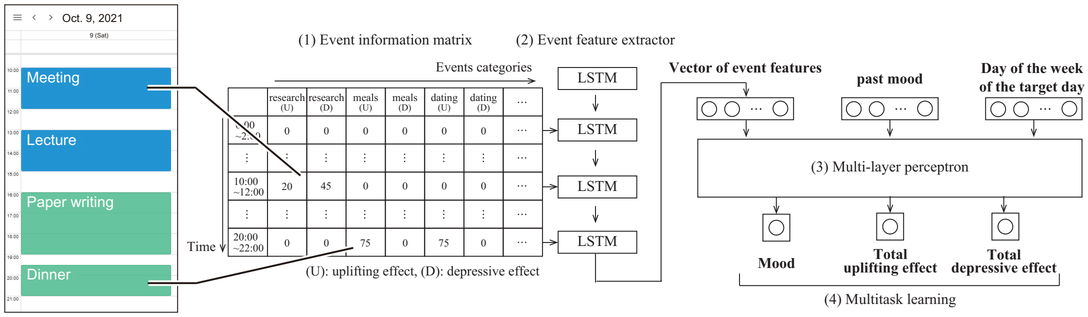

Fig. 2. Overview of the mood forecast model using daily events.

3.2 Forecast model

To forecast the one-week mental health indicators, we extended the mood prediction model using calendar events proposed by Tateyama et al. [48]. Despite the black-box nature of such a model, the model was reported to correlate the uplift of mood and positive events and the decline of mood and negative events. We believe that this is key to making the forecast understandable to users.

The model integrates daily events with historical mood data and day-of-the-week patterns to forecast moods, leveraging a 12×36 event matrix to capture daily event dynamics as shown in Figure 2. Rows represent 12-time steps separated by 2 hours from 0:00 to 24:00, and columns represent uplifting & depressive effects for each event category at each time step. Using each row as an input to long short-term memory (LSTM), the model can consider temporal patterns between events. A multilayer perceptron consolidates the output from LSTM with the past three days' mood trends and the day of the week information. Employing multitask learning, the model concurrently predicts mood and the aggregate total uplifting and depressive effects, balancing the influence of immediate events and historical mood data.

Model implementation and extension to one-week forecasting.
A preliminary experiment was conducted to collect calendar events and mood data to train the model. This study assumes a homogeneous group of participants, undergraduate and graduate students, which allows us to reduce the size of the dataset to train the model. A total of 20 undergraduate and graduate students (mean age: 24.1 ± 1.25) participated in the experiment. Participants were requested to record their daily events and mood information using the lifelog collection interface described in the previous section. Although the experiment period was not fixed, all participants completed data entry for at least two months. The data set collected in the preliminary experiment totaled 1324 days. Because there was missing event and mood information, only days with more than 6 hours of event registration were used, and if the mood was missing, it was supplemented with the previous day's value. The model was implemented to consider the past three days' mood. The number of neurons in the LSTM model was 32, and the middle layer of the multilayer perceptron was set to 16. The sigmoid function was used as the activation function. The results of training of the forecast model demonstrated that mood estimation accuracy was 0.0148 on the MSE when the range of mood values was 0 to 1. This prediction accuracy is comparable to previous studies that utilized events for forecasting mood [19, 45].

To extend the above prediction model to one-week forecasting, we used the future events and forecasted mood as inputs to the model. While the prediction model is adept at identifying patterns from historical data, it cannot predict future events. To bridge this gap, we leveraged the human ability to predict the potential effects of forthcoming events, a cognitive process that synthesizes experiential knowledge and contextual understanding [52]. In this study, we harnessed this human predictive capacity by incorporating anticipated effects of future events, as pre-assessed by users, into our model. Subsequently, the model's output, which is the forecasted mood for the immediate next day, was fed back into the model as a component of the input set for the subsequent day's forecasts. This recursive approach refines daily mood projections over a week.

3.3 Mental health indicator which is perceived to reflect past actions

Traditional mental health measures like Quick Inventory of Depressive Symptomatology (QIDS) [41] and Perceived Stress Scale (PSS) [13] require recollection of past feelings, posing a risk of recall bias. While these approaches are widely employed as reliable mental health indicators, because participants must recall information over a long period, recall bias [50] may introduce difficulty in capturing day-by-day behavior.

Our mental debt indicator is designed to be stable against daily fluctuations, focusing on the cumulative mental state. This indicator was inspired by the concept of profit and loss statements in accounting.

Let M_t denote the mood on a given day; the change from the previous day can be expressed as ΔM_t. Let P_t (P denotes profit) be the total uplifting effect and L_t (L denotes loss) be the total depressive effect, collected by the lifelog collection interface. Considering that mood is easily affected by a single event [15], it is assumed that ΔM_t is positively large when P_t is large and is negatively large when L_t is large. The large absolute value of ΔM_t makes it more straightforward to draw attention to local variations when viewed as a time series [11]. We defined daily change as ΔD_t using Equation (1). In this equation, ΔD_t is in the form of ΔM_t canceling with P_t and L_t out. Compared to ΔM_t, local fluctuations caused by events are expected to be suppressed in ΔD_t.

By adding previous ΔDs, a mental health indicator D can be calculated that reflects past situations. The method of addition and the period to be considered is important in the addition of ΔD. We adopted the weighted addition model that prioritizes recent states and does not consider the past beyond a certain period as shown in Equation (2). i denotes the number of days back from day t. In this equation, the ratio of ΔD added to D decreases as the past becomes more distant. Given the commonality of working during weekdays and taking a vacation during weekends, this study considers the past one week.

These definitions are inspired by accounting. Accounting involves "profit" which increases the "asset" and "loss" which decreases the "asset". An "asset" also includes "debt", which is borrowed from outside sources. The change in "debt" can be expressed using Equation (3). Comparing Equation (3) and Equation (1), we named the newly designed mental health indicator "mental debt". Participants can link high values with mental stress due to unaffordability and low values with ease of mental affordability. "Mental debt" aggregates mood data akin to financial trends, offering a long-term perspective that reflects past situations.

The forecasting model detailed in Section 3.2 can also be applied to forecast mental debt. The model also forecasts the total uplifting and depressive effects simultaneously. By using the forecasted mood and total uplifting/depressive effects, the forecast of mental debt can be calculated.

$$\Delta D_t := \Delta M_t - P_t + L_t \tag{1}$$

$$D_t := \sum_{i=0}^{6} \frac{7-i}{7} \Delta D_{t-i} \tag{2}$$

$$\Delta(\text{Debt}) = \Delta(\text{Asset}) - \text{Profit} + \text{Loss} \tag{3}$$

3.4 Reflection and planning application

The following functions were implemented as a PI system to utilize the forecast in a practical real-life setting.

- Self-reflection based on past events and moods.
- Future event planning concerning forecasted mental health indicators.

Fig. 3. Interfaces of the application for self-reflection and future planning. Text on the application was displayed in the local language. (A) The home view of the application. (B) The view for reflecting the participant's past. (C) The view for registering future events. (D) The view for planning future events.

Self-reflection based on past events and moods. The home screen, shown in Figure 3 (A) features a month's worth of mental health indicators and a weekly calendar with events recorded by participants and Google Calendar API-retrieved events. Clicking on the graph displays the daily average hours for event categories in a bar-graph style as shown in Figure 3 (B). In addition, by selecting a different day or period, the daily average hours can be compared across multiple periods. Participants can focus on the characteristic change trends and points in the mental health indicators and analyze how their time has been spent. This assists in identifying positive and negative ways of spending time.

Future event planning concerning forecasted mental health indicators. "Today" is marked on the graph (Figure 3 (A)) with a red line, with forecasts to the right. Participants input future events via the interface in Figure 3 (C) and assess their expected impact. This interface can be opened by tapping the icon in the upper right corner of the home screen. Here, a list of future events for next week is displayed. The app displays a question as follows: "How much impact do you think this will have?". The system also accommodates habitual events like work not listed in the calendar. Participants select the time of the event and register it as "How I spend my weekdays".

Events can be planned either by specifying a particular time directly on the calendar, or by tapping the button in Figure 3 (A). Participants can plan events by referring to the forecasted values displayed on the graph using the interface shown in Figure 3 (D). The forecasted values are updated in real time according to the input of the effects of the planned event.

4 EXPERIMENT

4.1 Field study

The experiment aimed to monitor the cognitive and behavioral responses to the forecast regarding research questions. Experimental procedures for this study were approved by the appropriate ethics review committee and informed consent was obtained prior to participation. For participant recruitment, an email announcement was sent to the author's university and an announcement was posted on Twitter, targeting undergraduate and graduate students. 133 students registered for the experiment using an application form. The details of the experiment were explained at the time of recruitment, and compensation of approximately USD 50 was offered for the completion of the experiment. The experimental period was approximately four weeks and was considered complete when both the daily and weekly tasks were completed by the fifth Sunday after the start of the experiment.

***Daily tasks*.**

This task was designed to collect information that enables the system to forecast future mental health indicators and allows participants to reflect on their lives in the application. Participants were requested to record their events each day throughout the experiment. Every day at a fixed time (8:30 or 9:00 PM. by default), an email with the URL for the application was sent to prompt participants to enter their event information and mood. In the application, the lifelog collection interface shown in Section 3.1 was displayed. After answering the questions, participants were subsequently redirected to the home screen shown in Figure 3(A). Participants were free to reflect on past events and plan future events. To reduce recall bias, participants were only allowed to record up to one previous day's events and moods.

***Weekly tasks*.**

This task was designed to encourage participants to plan future events on the application. Because event planning using the application was not a specifically required step for participants, this feature may not be used. Therefore, participants were requested to perform event planning every Sunday evening. After recording their event information and mood, participants received a URL to a separate web-based form and were instructed to complete the following instructions.

(1) Record the category and impact of each event over the following week from Figure 3 (C).
(2) Plan future events over the following week from Figure 3 (D).

To ensure that participants plan work time as well as leisure time, two questions "Did you plan an uplifting event?" and "Did you plan a depressive event?" were asked.

After completing the planning of future events, participants were asked to respond to the perceived stress scale (PSS) [13]. The PSS is an indicator consisting of 10 questions that take values from 0 to 40, with higher values indicating that individuals are more stressed. Participants were requested to consider the events of the past week.

***Experimental settings*.**

The experiment was designed with the primary aim of exploring two research questions (RQs). The first, RQ1, investigated how a one-week forecast affects participants' thoughts and behaviors. To prove this, we created two distinct intra-participant conditions: a forecast-present condition and a forecast-absent condition. Participants were initially exposed to one of these conditions, with a switch occurring on the third Sunday—thereafter, those who had seen the forecast would not, and vice versa.

Addressing the second question, RQ2, required an examination of how participants engaged with different mental health indicators. To avoid confusion that might arise from the concurrent use of multiple indicators, we introduced inter-participant conditions by dividing participants into two groups: one that focused on mood and another on mental debt. For the mental debt group, a detailed explanation was provided through a pop-up window, which appeared after daily tasks and could also be accessed anytime by tapping the "Meaning of Indicators" button (as illustrated in Figure 3(A)).

- The “mental debt” is calculated by considering not only the day’s mood but also the impact of the events of the previous seven days.
- Higher values (upper side) indicate that there were more depressive effects on the mood of events.
- Lower values (lower side) indicate that there were more uplifting effects on the mood of events.

To ensure the rigor of the study, the sequence of conditions — both intra-participant (forecast timing) and inter-participant (indicator type) — was randomized at the time of participant registration.

4.2 Qualitative analysis

Based on the evaluation framework for the HCI system addressing behavior change [23], we conducted a qualitative analysis aiming to reveal why and how users had engaged with the forecast and its indicators, offering insights into their interaction with the system. An open-ended questionnaire regarding the forecast was conducted after the fifth weekly task. Participants were requested to respond to the following questions: “Describe the difference in your thoughts and actions between the period when the forecast was displayed and the period when it was not displayed.” After the fifth weekly task, participants were invited to participate in a voluntary interview. The interview was conducted in a semi-structured format with the question, mainly about “Was there any impact of the forecast?” We asked to think about the differences between the period with and without the forecast. In the interview, participants also think about how they reflect on their events using the application and how they think about their mental health situations. The interview lasted about half an hour, enough time to hear about the experimental experience.

A total of 105 participants (male: 48, female: 57, age: 23.6 ± 2.52) recorded their events and moods and completed weekly tasks for four weeks. The number of participants in the mood group was 52 and 53 in the mental debt group. 30 participants took part in the interviews (15 for the mood group and 15 for the mental debt group). Out of the 105 participants who completed the 4-week experiment with mood logs and weekly tasks, six did not respond to the questionnaire and interview. The results of both the questionnaires and interviews represent the feedback collected from 99 participants.

Three authors of this study participated in the qualitative analysis. The three authors read the transcribed records of the questionnaires and interviews and understood the contents. In the coding process, open coding was performed and codes were created to classify references to the forecast. The three authors discussed the codes and decided on the final categories. The codes were broadly divided into two categories: “effects of a one-week forecast of mental health indicators” and “factors preventing the usage of the forecast.” Next, activity coding was performed and the categories were assigned to each mention of the forecast. We discussed the categories that two or more authors agreed on as the final assignment. We calculated the number of references for each category. If the same category was mentioned in both the questionnaire and the interview, it was counted only once.

5 RESULTS

In this section, we investigate the data procured from our experiment to evaluate the effects of forecasts and to understand the underlying causes of its outcomes. We begin by presenting quantitative data that validates the application of our intervention. Subsequently, we provide an in-depth qualitative analysis that offers insights into participants’ responses and their perceptions of the forecasts. We then address our primary research questions in two dedicated sections, integrating both the quantitative measurements and qualitative findings.

Table 2. The ratio of Task, Activity, Life, and Other events to the total number of events that participants experienced and planned. The upper part shows the ratio of events that participants experienced, and the lower part shows the ratio of events that participants planned. The sum of the ratio exceeds 1 because multiple labels can be assigned to an event.

	Task	Activity	Life	Other
Experienced events				
preliminary experiment	0.63	0.25	0.14	0.045
this experiment	0.59	0.22	0.28	0.015
- (mood group)	0.59	0.22	0.29	0.025
- (mental debt group)	0.59	0.21	0.26	0.0044
Planned events				
mood group (without forecast)	0.56	0.22	0.21	0.013
mental debt (without forecast)	0.56	0.14	0.17	0.0099
mood group (with forecast)	0.57	0.18	0.20	0.15
mental debt (with forecast)	0.54	0.17	0.18	0.13

5.1 Behavioral patterns of participants

To confirm the applicability of the forecast model developed in the preliminary experiment, we evaluated the consistency between its data and the current study's data. The Bartlett test confirmed that the variances of mood distributions were equal ($p = 0.285$). The unpaired t-test confirmed that there was no significant difference in moods between the preliminary experiment and this study ($p = 0.104$), suggesting comparability of the two mood datasets.

We also focused on how participants spent their days. We compared the ratio of Task, Activity, Life, and Other events based on the classification of event labels in Table 1. The upper part shows the ratio of events that participants experienced, and the lower part shows the ratio of events that participants planned on the application. There seems to be no specific difference in the ratio of experienced events between the preliminary experiment and this study and between the mood group and the mental debt group. This similarity suggests a lifestyle homogeneity between groups, validating the use of the same forecast model and comparing the effects of the forecast between these groups.

We also investigated planned events. Figure 4 shows the number of future events registered on the application with and without the forecast. Results were found to be non-normally distributed according to the Shapiro-Wilk test at the 5% level. The Mann-Whitney U test showed that there were no significant differences between with and without the forecast in both mood ($p = 0.455$) and mental debt groups ($p = 0.732$). The ratio of events that participants planned is shown in the lower part of Table 2. This suggests that there were no significant differences between mood and mental debt groups in the pattern of planning.

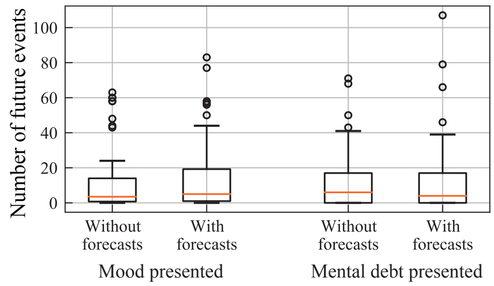

Fig. 4. Number of future events registered on the application.

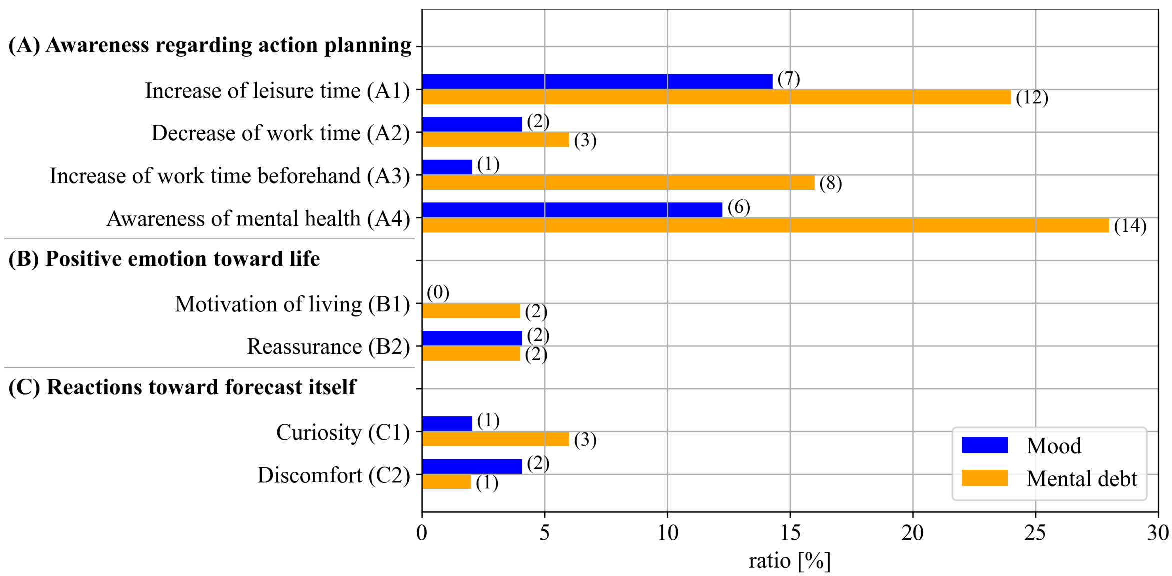

Fig. 5. The count of codes regarding the effects of a one-week forecast of mental health indicators. The numbers at the tips of the bars indicate the number of respondents.

5.2 Overview of qualitative analysis

*5.2.1 **Effects of a one-week forecast of mental health indicators**.*

Figure 5 classifies the codes regarding the effects of the forecasted mental health indicators. Overall, 32 participants in the mood group and 17 participants in the mental debt group did not report the effects of the forecast. Figure 5 reflects participants who reported the effects of the forecast.

(A) Awareness regarding action planning.

The forecast led to changes in participants' awareness to improve their mental health.

Increase of leisure time (A1) indicates that the forecast promoted to take a break, and **Decrease of work time (A2)** indicates that the forecast promoted to reduce workload.

Increase of work time beforehand (A3) indicates that the forecast made participants do the workload in advance. This item was related to the awareness of reducing future workload, as mentioned, "*When the forecast* *was displayed, I tried to do something difficult when the workload was small* (No. 154, mental debt group)." Only one participant mentioned this in the mood group, whereas eight made this assertion in the mental debt group.

Awareness of mental health (A4) indicates that the forecast increased awareness of their mental health considerations and action plans. This item was included in "*I paid more attention to my mood* (No. 104, mood group)." which directly refers to mental health.

The above items are more reported by the mental debt group than the mood group, which means the forecast of mental debt is more effective for awareness of action planning.

***(B) Positive emotion toward life*.**

Some participants felt positive emotions toward their lives after seeing the forecast.

Motivation of living (B1) indicates that participants looked forward to spending the forecasted days, as one expressing, "*In general, I was looking forward to tomorrow when there was a forecast* (No. 108, mental debt group)."

Reassurance (B2) indicates that the forecast provided a sense of security because it assists participants in anticipating their future. This item includes "*Without the forecast, I feel anxious because I don't know what will happen in the future* (No. 55, mental debt group)."

***(C) Reactions toward forecast itself*.**

Some participants reacted to the forecast itself.

Curiosity (C1) indicates that participants were curious about the forecasted values, as one expressing, "*I've never seen a forecast of a subjective indicator like mental health, so I thought it was quite interesting* (No. 49, mental debt group)." This reaction can be a factor in enhancing the utilization of the forecast.

In contrast, **Discomfort (C2)** was observed. This indicates that participants experienced discomfort with the forecast. One such comment was, "*If I start thinking that it's going to be tough, it often indeed becomes tough* (No. 127, mood group)." This reaction was caused by the pessimistic thinking of the participants.

5.2.2 Factors preventing the usage of the forecast.

Figure 6 classifies the codes regarding factors that prevented the usage of the forecast.

***(X) System-specific factors*.**

Some factors preventing the usage of the forecast were specific to this study's system. These factors prevented participants from feeling that the forecast was accurate.

Low trust in the algorithm (X1) indicates algorithmic doubt regarding the forecast model. As one expressed "*I felt that the forecast was far from what I thought* (No. 97, mental debt group)," dissatisfaction with the forecast's accuracy prevented the use of the forecast.

Important information except for the forecast (X2) indicates that some participants emphasized factors other than the forecast. Participants mentioned sleep time, menstrual cycle, and autonomic imbalance as factors other than events that affect their mental health. Because the forecast does not consider these factors, participants did not use the forecast.

Lack of inputs for the forecast (X3) indicates that participants did not input enough information to the system. As expressed, "*I felt a little bothered to enter future events* (No. 88, mood group)," and "*I didn't plan events because I had never used a calendar app* (No. 55, mental debt group)," the forecast model could not make an appropriate forecast because the necessary information was not inputted. Participants did not use the forecast because they recognized that the forecast model functioned incorrectly.

Misunderstanding of the forecast mechanism (X4) indicates that participants did not realize that the forecast model considered future events. As expressed, "*I thought the forecast was based on past events rather than future events* (No. 80, mood group)," participants did not realize that planning for the future would lead to improved forecasts in the future.

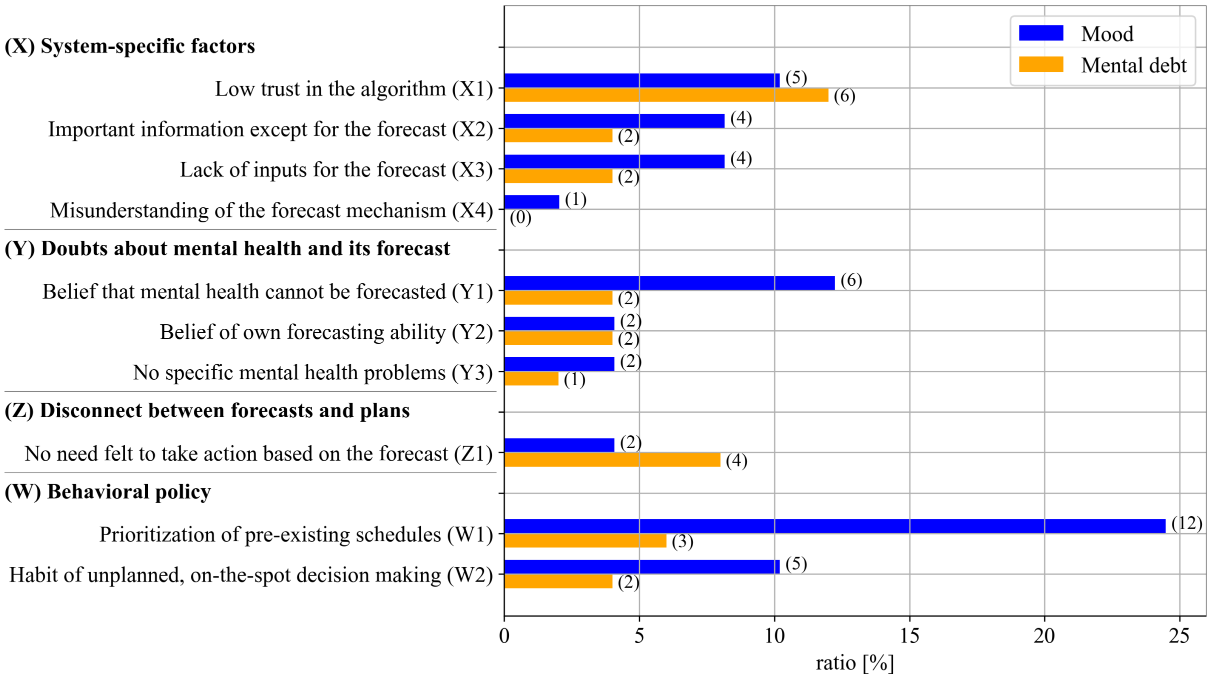

Fig. 6. The count of codes regarding factors preventing the usage of the forecast. The numbers at the tips of the bars indicate the number of respondents.

(Y) Doubts about mental health and its forecast.

Doubts about mental health and its forecast also prevented the use of the forecast.

Belief that mental health cannot be forecasted (Y1) indicates that participants felt forecasting itself was impossible because of unexpected future events. As one expressed "*mood is determined by whether there is a good event or a bad event in the future* (No. 102, mood group)," some participants thought mental health was largely influenced by factors on the day of the event. This item was more frequently reported in the mood group.

Belief of own forecasting ability (Y2) indicates that participants could assume their mental health in their own minds. As one expressed "*I can forecast my mental health for the next few days because I know my schedule* (No. 148, mental debt group)," some participants did not feel the necessity to use the forecast because they could forecast their mental health by themselves.

No specific mental health problems (Y3) indicates that participants did not feel the necessity to care own mental health. As one expressed "*I'm not in mentally ill, so I'm okay* (No. 55, mental debt group)," some participants did not feel the necessity to use the forecast because they did not have any mental health problems.

(Z) Disconnect between forecasts and plans.

The forecast did not always recall action planning.

No need felt to take action based on the forecast (Z1) indicates that participants were not motivated to take action based on the forecast. As one expressed "*The forecast was only increasing, not decreasing, so I thought it was okay. If it had been decreasing, I might have thought about it* (No. 102, mood group)," the forecast was not always informative for participants to change their actions.

Table 3. Changes in PSS for each experimental condition (mean and standard deviation)

Indicator / Timing of forecast	Week 0	Week 1	Week 2	Week 3	Week 4
Mood / First (n = 23)	23.3 ± 6.60	20.8 ± 7.16	20.0 ± 6.55	20.3 ± 6.17	22.7 ± 7.84
Mood / Second (n = 29)	19.5 ± 6.12	17.6 ± 7.72	20.2 ± 8.81	18.5 ± 7.71	20.5 ± 8.93
Mental debt / First (n = 21)	22.1 ± 5.64	17.8 ± 6.12	18.9 ± 4.80	19.0 ± 6.47	19.8 ± 5.80
Mental debt / Second (n = 32)	21.3 ± 7.40	20.5 ± 8.52	19.4 ± 7.73	19.8 ± 7.96	19.8 ± 6.99

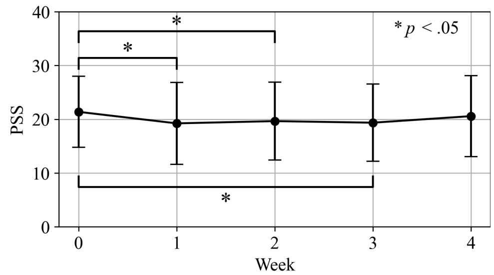

Fig. 7. Changes of means of PSS. The error bars indicate the standard error.

(W) Behavioral policy.

Some participants did not use the forecast because of their behavioral policy.

Prioritization of pre-existing schedules (W1) indicates that participants prioritized to follow their own schedules. As one expressed "*I had a lot of plans that were decided externally, so there was no difference* (No. 148, mental debt group)," some participants felt it was difficult to change their schedules. This tendency was more frequently reported in the mood group.

Habit of unplanned, on-the-spot decision making (W2) indicates that participants did not always plan their events. As one expressed, "*I think about my work on weekdays and prioritize it, and I want to reduce the time I spend planning* (No. 99, mental debt group)," some participants did not like to plan in advance.

5.3 How is a one-week forecast of mental health indicators recognized and utilized? (RQ1)

Table 3 shows the mean and standard deviation of the changes in the perceived stress scale (PSS), with higher values indicating more stress, under each experimental condition. "Week" refers to the week after the start of the experiment, and "Week 0" refers to the start of the experiment. The results excluding only "Week 4" of Mood / Second were found to be normally distributed according to the Shapiro-Wilk test at the 5% level. Three-factor analysis of variance was conducted, with the type of indicator as displayed to participants in the web application (mood and mental debt group) and the timing when the forecast was displayed (first and second half of the experimental period) as the inter-group factors, and repeated measures for each participant as the intra-group factors. The analysis of variance revealed only one main effect of the repeated measures factor ($F(4, 505) = 4.20, p = 0.0036, \eta^2 = 0.0139$). No first or second-order interactions were identified. Figure 7 shows the change in the mean of PSS. Multiple comparisons adjusted for p values by Shaffer's method showed that PSS was significantly lower during weeks 1 (p = 0.0105), 2 (p = 0.0105), and 3 (p = 0.022) compared to those at "Week 0". There were no significant differences in PSS between "Week 0" and "Week 4" or between the other conditions.

Focusing on the qualitative analysis, revealed how participants recognized and utilized the one-week forecast. Regarding the effects of the forecast, participants reported that they utilized the forecast to change their actions.

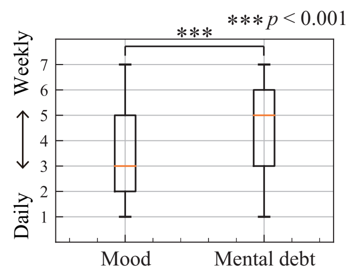

Fig. 8. Perceived target period using the indicators.

They utilized the forecast to increase their leisure (A1) time and reduce their work time (A2). Increased awareness of mental health (A4) was also reported. While previous studies also focused on the awareness of improving mental health, this study found that an "increase of work time beforehand (A3)" occurred. A longer forecast period broadens the range of coping strategies beyond simply taking leisure time. Participants recognized their tendency to mental health and changed actions to reduce their future workload.

Apart from practical action strategies, the one-week forecast also made participants feel positive emotions toward their lives. "Motivation of living (B1)" and "reassurance (B2)" indicate that participants can positively perceive their future. One-week forecast was recognized as a barometer of future conditions, which brought positive feelings.

The forecast itself can be both a positive and negative factor for utilization. Some participants recognized the one-week forecast as a curious thing (C1), while others recognized it as an uncomfortable thing (C2).

5.4 If a forecast of mental health indicators can reflect past actions, how does the forecast affect behaviors? (RQ2)

Mood can be easily interpreted because it is subjectively answered. However, it is not clear how mental debt is interpreted. To examine how participants perceived mental debt, the relationship between mental debt and PSS was analyzed. Because PSS varies among individuals depending on their perceptibility of stress and other factors, a linear mixed model was used for analysis. Mental debt D for each participant can be expressed as Equation (4) using PSS. i is the index of each data set and j represents the index of each participant. β_0 and β_1 denote the fixed effects of the intercept and slope common among participants, respectively, and $\beta_{0,j}$ and $\beta_{1,j}$ denote the random effects of the intercept and slope for each participant, respectively.

$$D_{i,j} = \bar{\beta}_{0,j} + \bar{\beta}_{1,j}\mathrm{PSS}_{i,j} + \epsilon_{i,j} \tag{4}$$
$$(\bar{\beta}_{0,j} = \beta_0 + \beta_{0,j}, \bar{\beta}_{1,j} = \beta_1 + \beta_{1,j})$$

The fixed effects of the linear mixed model were $\beta_0 = -119.7(p < 0.001)$ and $\beta_1 = 3.551(p < 0.001)$. The slope of β_1 with respect to PSS was significantly positive, which indicates that mental debt correlates to PSS.

Figure 8 summarizes how participants used mental health indicators to consider a day or a week. On the Likert scale, "1" corresponds to "Think about how to spend one **day**" and "7" for "Think about how to spend one **week**." The mental debt group utilized the indicator more to "think about how to spend one **week**" than the mood group. The results were found to be non-normally distributed according to the Shapiro-Wilk test at the 5% level. The Mann-Whitney U test showed that the median value was significantly higher for mental debt ($p < 0.001$). These results suggest that mental debt was perceived to reflect long-term past situations.

Based on the qualitative results, some differences were identified between mood and mental debt. Focusing on the overall tendency, mental debt had more impacts on behavior than mood. While "increase of leisure time (A1)", "awareness of mental health (A4)", and "increase of work time beforehand (A3)" mainly occurred in the mental debt group. The case of "decrease of work time (A2)" was less popular in both groups. "Increase of work time beforehand (A3)" was the most differentiating effect between the two groups. Only one participant mentioned this in the mood group, whereas eight made this assertion in the mental debt group. This suggests that mental debt which reflects past situations affects more effectively taking actions in advance.

Regarding the factors that prevented the usage of the forecast, there were differences in trends between the mood and mental debt groups. "Belief that mental health cannot be forecasted (Y1)" was a characteristic inhibitory factor. Six participants mentioned this in the mood group, whereas two pointed it out in the mental debt group. This suggests that mental debt has the effect of mitigating the perception that mental health is affected by unpredictable factors.

The mood group tended to follow their behavioral policy, whereas the mental debt group did not. The most differentiating item between mood and mental debt was "prioritization of pre-existing schedules (W1)." This was mentioned by 12 participants in the mood group, whereas it was mentioned by only three in the mental debt group. "Habit of unplanned, on-the-spot decision making (W2)" was also mentioned by five participants in the mood group and only two in the mental debt group. This suggests that the forecast reflecting past situations is less affected by inhibitory factors that prevent the consideration of actions in advance.

6 DISCUSSION

6.1 Longer perspective with a one-week forecast

The application showed improvement in participants' mental health as per the perceived stress scale, with higher values indicating more stress, though not directly attributable to the one-week forecast feature. The number of planned events did not differ significantly between the mood and mental debt groups, possibly because event planning was the weekly task, regardless of whether the forecast was presented. Under both conditions, opportunities were provided for participants to consider their future and improvement. It is necessary to verify in the future whether the forecast increases event planning based on spontaneous planning only.

According to Figure 5, similar to previous studies [19], the forecast affected participants in planning rest time and working less and awareness of mental health. However, the "increase of work time beforehand (A3)" is a characteristic result of the one-week forecast. This type of planning strategy may be more feasible because it provides practical support for a reduction in future burdens. We believe that this occurred because the one-week forecast enabled participants to consider situations other than the present, whereas, in the case of a short-period forecast, planning is mostly to improve the situation close to the present.

A longer perspective also produced both positive and negative emotions. The farther into the future, the greater the difficulty in envisioning one's mental state [16]. A longer forecast period has the potential to assist in addressing this challenge and reduce stress caused by future uncertainty [39]. Furthermore, even if the forecasted state is negative, it may lead to positive actions to reduce it [38]. These perceptions of the future lead to enhancing the motivation of living and reassurance. However, some felt discomfort by the negative forecast. Negative forecasts may reduce self-efficacy because they suggest that one cannot cope with future events [4]. For those who are pessimistic or vulnerable to stress, the forecast may increase anxiety. In addition, excessive avoidance of negative events may harm daily life and social relationships. To address this, it is essential to offer remedial strategies beyond simply toggling the forecast display. Future enhancements to our system could include coping strategies for anticipated stressors and encourage collaboration with cognitive behavioral therapy, such as cognitive reappraisal [17], helping users constructively engage with future anxieties, promoting resilience and proactive mental health management.

6.2 A sense of the connection between the past and the forecast

Mental debt was found to bring a sense of the connection between the past and the forecast. The qualitative analysis results suggest that this sense is an important factor in designing systems that utilize forecasts of subjective indicators for behavioral change.

One of the effects of this sense is strategic action planning. "(A) Awareness regarding action planning" was more frequent in the mental debt group, especially in "increase of work time beforehand (A3)." Mental debt enabled participants to recognize that a series of actions would affect the future, which encouraged them to take action to improve their future state. It introduced a view of a balance between leisure and work time, rather than simply planning leisure time. This view is important for long-term self-management because it promotes problem-focused coping as well as emotion-focused coping.

A sense of the connection between the past and the forecast also affected the value of action planning based on the forecast. "(W) Behavioral policy" and "belief that mental health cannot be forecasted (Y1)" were reported especially in the mood group. The forecast of mood is easily fluctuated by a single event. Some participants indicated that "belief that mental health cannot be forecasted (Y1)", for example, "My mood here depends on whether the day's experiments went well or not (No. 98, mood group)." When viewing the future from a local perspective, it is easy to be conscious of time constraints, which resulted in "Prioritization of pre-existing schedules (W1)." For those who have a "Habit of unplanned, on-the-spot decision making (W2)," the forecast which easily fluctuates by a single event cannot provide an incentive to plan beforehand. In contrast, mental debt reflects past situations and gradually changes. Therefore, it broadens participants' view of planning, and it is easier to find room for coping. Users can recognize that past actions are linked to future states, which makes them recognize the value of proactive actions.

6.3 Limitations

There are limitations in the system implementation in real life. The study's participant base, consisting of undergraduate and graduate students, may not accurately represent the responses of a broader population, such as full-time workers. Consequently, further research is needed to assess the system's effectiveness across diverse demographics and lifestyles.

The forecast period was not sufficiently long to enable participants to change schedules. Some participants indicated that they could not change their schedules one week in advance because they had routine work. Future iterations could explore extending the forecast period to better accommodate such users.

Another challenge is the user burden of inputting comprehensive data for accurate forecasting. Missing or incomplete data can undermine forecast reliability, leading to decreased user trust and engagement. Simplifying data entry and enhancing the system's ability to generate accurate forecasts with limited input could improve user experience and forecast utilization.

Lastly, not all participants found the application supportive in managing their mental health, suggesting a disconnect between the forecast and its practical use. Future enhancements might include actionable suggestions tied to forecasted mental states, such as recommending specific leisure activities, to provide more tangible planning assistance.

7 CONCLUSION

In this study, we presented a one-week forecast of mental health indicators and investigated how the forecast was recognized and utilized. Additionally, we investigated how the perception of reflecting the past situations affects behaviors. We developed a novel indicator, described as "mental debt," which better reflects past circumstances than mood. In the experiment, participants recorded their events and moods for four weeks, utilizing a web application for self-reflection and event planning. The effects of a one-week forecast of mental health indicators and factors preventing the usage of the forecast were categorized. In particular, "increase of work time beforehand (A3)" was identified as an important effect of "(A) awareness regarding action planning." One-week forecast was also recognized as a barometer of future conditions, which brought positive feelings. One-week forecast provided a longer perspective than daily mood, which enabled participants to foresee the future.

Mental debt was more likely to be perceived as reflecting past situations. The mental debt group was more likely to lead to an "increase of work time beforehand (A3)" which is a strategic action planning. In contrast, the mood group was more likely to report factors that prevented the utilization of the forecast. The mood was perceived as something affected by unpredictable factors. In mood presented group, behavioral policies such as prioritization of pre-existing schedules were more likely to be reported. These results suggest that a sense of the connection between the past and the forecast is an important factor in designing systems that utilize forecasts of subjective indicators for behavioral change.

In future research, we aim to leverage the forecast to foster behavioral change. This will involve examining whether a process of pointing out specific days that must be dealt with based on the forecast, will promote future planning. We will also investigate whether presenting a concrete plan of feasible events will promote forecast utilization.

ACKNOWLEDGMENTS

This work was supported by JSPS KAKENHI under Grant JP19J22452.

REFERENCES

[1] Ane Alberdi, Asier Aztiria, and Adrian Basarab. 2016. Towards an automatic early stress recognition system for office environments based on multimodal measurements: A review. *Journal of Biomedical Informatics* 59 (2016), 49–75. https://doi.org/10.1016/j.jbi.2015.11.007

[2] Deemah Alqahtani, Caroline Jay, and Markel Vigo. 2020. The role of uncertainty as a facilitator to reflection in self-tracking. *DIS 2020 - Proceedings of the 2020 ACM Designing Interactive Systems Conference*, 1807–1818. https://doi.org/10.1145/3357236.3395448

[3] Amna Amanat, Muhammad Rizwan, Abdul Rehman Javed, Maha Abdelhaq, Raed Alsaqour, Sharnil Pandya, and Mueen Uddin. 2022. Deep Learning for Depression Detection from Textual Data. (2022), 1–13.

[4] Albert Bandura. 1982. Self-efficacy mechanism in human agency. *American Psychologist* 37, 2 (feb 1982), 122–147. https://doi.org/10.1037/0003-066x.37.2.122

[5] Albert Bandura. 2001. Social Cognitive Theory: An Agentic Perspective. *Annual Review of Psychology* 52, 1 (2001), 1–26.

[6] Vicky Bilbily, Elaine Huynh, Karan Singh, and Fanny Chevalier. 2021. Space, Time, and Choice: A Unified Approach to Flexible Personal Scheduling. In *The 34th Annual ACM Symposium on User Interface Software and Technology* (Virtual Event, USA) *(UIST '21)*. Association for Computing Machinery, New York, NY, USA, 484–497. https://doi.org/10.1145/3472749.3474764

[7] Ilona Boniwell and Philip G Zimbardo. 2004. Balancing Time Perspective in Pursuit of Optimal Functioning. *Positive psychology in practice.* (2004), 165–178.

[8] Bruce K Britton and Abraham Tesser. 1991. Effects of Time-Management Practices on College Grades. , 405-410 pages. Issue 3.

[9] Aaron Frederick Bulagang, Ng Giap Weng, James Mountstephens, and Jason Teo. 2020. A review of recent approaches for emotion classification using electrocardiography and electrodermography signals. *Informatics in Medicine Unlocked* 20 (2020), 100363. https://doi.org/10.1016/j.imu.2020.100363

[10] Prerna Chikersal, Afsaneh Doryab, Michael Tumminia, Daniella K. Villalba, Janine M. Dutcher, Xinwen Liu, Sheldon Cohen, Kasey G. Creswell, Jennifer Mankoff, J. David Creswell, Mayank Goel, and Anind K. Dey. 2021. Detecting depression and predicting its onset using longitudinal symptoms captured by passive sensing: A machine learning approach with robust feature selection. *ACM Transactions on Computer-Human Interaction* 28, 1 (2021). https://doi.org/10.1145/3422821

[11] Eun Kyoung Choe, Bongshin Lee, Haining Zhu, Nathalie Henry Riche, and Dominikus Baur. 2017. Understanding self-reflection: How people reflect on personal data through visual data exploration. *PervasiveHealth: Pervasive Computing Technologies for Healthcare* (2017), 173–182. https://doi.org/10.1145/3154862.3154881

[12] Brigitte J.C. Claessens, Wendelien Van Eerde, Christel G. Rutte, and Robert A. Roe. 2007. A review of the time management literature. *Personnel Review* 36 (2007), 255–276. Issue 2. https://doi.org/10.1108/00483480710726136

[13] S. Cohen, T. Kamarck, and R. Mermelstein. 1983. A global measure of perceived stress. *Journal of health and social behavior* 24, 4 (1983), 385–396. https://doi.org/10.2307/2136404

[14] Daniel A. Epstein, Clara Caldeira, Mayara Costa Figueiredo, Xi Lu, Lucas M. Silva, Lucretia Williams, Jong Ho Lee, Qingyang Li, Simran Ahuja, Qiuer Chen, Payam Dowlatyari, Craig Hilby, Sazeda Sultana, Elizabeth V. Eikey, and Yunan Chen. 2020. Mapping and Taking

Stock of the Personal Informatics Literature. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 4, 4, Article 126 (dec 2020), 38 pages. https://doi.org/10.1145/3432231

[15] Julie A. Fuller, Gwenith G. Fisher, Jeffrey M. Stanton, Christiane Spitzmüller, Steven S. Russell, and Patricia C. Smith. 2003. A Lengthy Look at the Daily Grind: Time Series Analysis of Events, Mood, Stress, and Satisfaction. *Journal of Applied Psychology* 88, 6 (2003), 1019–1033. https://doi.org/10.1037/0021-9010.88.6.1019

[16] Daniel T Gilbert, Elizabeth C Pinel, Timothy D Wilson, Stephen J Blumberg, and Thalia P Wheatley. 2002. *Durability Bias in Affective Forecasting.* Cambridge University Press, 292–312. https://doi.org/10.1017/CBO9780511808098.018

[17] James J Gross. 2002. Emotion regulation: affective, cognitive, and social consequences. *Psychophysiology* 39, 3 (2002), 281–291. https://doi.org/10.1017/s0048577201393198

[18] Emily Hewlett and Valerie Moran. 2014. *Making Mental Health Count.* OECD Publishing. 244 pages. https://doi.org/10.1787/9789264250550-ko

[19] Victoria Hollis, Artie Konrad, Aaron Springer, Matthew Antoun, Christopher Antoun, Rob Martin, and Steve Whittaker. 2017. What Does All This Data Mean for My Future Mood? Actionable Analytics and Targeted Reflection for Emotional Well-Being. *Human-Computer Interaction* 32, 5-6 (2017), 208–267. https://doi.org/10.1080/07370024.2016.1277724

[20] Victoria Hollis, Alon Pekurovsky, Eunika Wu, and Steve Whittaker. 2018. On Being Told How We Feel: How Algorithmic Sensor Feedback Influences Emotion Perception. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2 (9 2018), 1–31. Issue 3. https://doi.org/10.1145/3264924

[21] Daniel Kahneman, Alan B Krueger, David A Schkade, Norbert Schwarz, and Arthur A Stone. 2004. A Survey Method for Characterizing Daily Life Experience: The Day Reconstruction Method. *Science* 306, 5702 (2004), 1776–1780. https://doi.org/10.1126/science.1103572

[22] Taewan Kim, Haesoo Kim, Ha Yeon Lee, Hwarang Goh, Shakhboz Abdigapporov, Mingon Jeong, Hyunsung Cho, Kyungsik Han, Youngtae Noh, Sung-ju Lee, and Hwajung Hong. 2022. Prediction for Retrospection: Integrating Algorithmic Stress Prediction into Personal Informatics Systems for College Students' Mental Health. In *CHI Conference on Human Factors in Computing Systems.* ACM, New York, NY, USA, 1–20. https://doi.org/10.1145/3491102.3517701

[23] Predrag Klasnja, Sunny Consolvo, and Wanda Pratt. 2011. How to Evaluate Technologies for Health Behavior Change in HCI Research. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada) *(CHI '11).* Association for Computing Machinery, New York, NY, USA, 3063–3072. https://doi.org/10.1145/1978942.1979396

[24] Todd Kulesza, Margaret Burnett, Weng Keen Wong, and Simone Stumpf. 2015. Principles of Explanatory Debugging to personalize interactive machine learning. *International Conference on Intelligent User Interfaces, Proceedings IUI* 2015-January, 126–137. https://doi.org/10.1145/2678025.2701399

[25] Richard S Lazarus and Susan Folkman. 1984. *Stress, appraisal, and coping.* Springer publishing company.

[26] Jungeun Lee, Sungnam Kim, Minki Cheon, Hyojin Ju, JaeEun Lee, and Inseok Hwang. 2022. SleepGuru: Personalized Sleep Planning System for Real-Life Actionability and Negotiability. In *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology* (Bend, OR, USA) *(UIST '22).* Association for Computing Machinery, New York, NY, USA, Article 52, 16 pages. https://doi.org/10.1145/3526113.3545709

[27] Kwangyoung Lee, Hyewon Cho, Kobiljon Toshnazarov, Nematjon Narziev, So Young Rhim, Kyungsik Han, Youngtae Noh, and Hwajung Hong. 2020. Toward Future-Centric Personal Informatics: Expecting Stressful Events and Preparing Personalized Interventions in Stress Management. (2020). https://doi.org/10.1145/3313831.3376475

[28] Kwangyoung Lee and Hwajung Hong. 2018. MindNavigator: Exploring the stress and self-interventions for mental wellness. *Conference on Human Factors in Computing Systems - Proceedings* 2018-April (2018), 1–14. https://doi.org/10.1145/3173574.3174146

[29] Boning Li and Akane Sano. 2020. Extraction and Interpretation of Deep Autoencoder-Based Temporal Features from Wearables for Forecasting Personalized Mood, Health, and Stress. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 4, 2, Article 49 (jun 2020), 26 pages. https://doi.org/10.1145/3397318

[30] Ian Li, Anind Dey, and Jodi Forlizzi. 2010. A stage-based model of personal informatics systems. *Conference on Human Factors in Computing Systems - Proceedings* 1 (2010), 557–566. https://doi.org/10.1145/1753326.1753409

[31] Ningyun Li, Huijun Zhang, and Ling Feng. 2021. Incorporating Forthcoming Events and Personality Traits in Social Media based Stress Prediction. *IEEE Transactions on Affective Computing* (2021), 1. https://doi.org/10.1109/TAFFC.2021.3076294

[32] Qi Li, Liang Zhao, Yuanyuan Xue, Li Jin, Mostafa Alli, and Ling Feng. 2017. Correlating Stressor Events for Social Network Based Adolescent Stress Prediction. In *Database Systems for Advanced Applications*, Selçuk Candan, Lei Chen, Torben Bach Pedersen, Lijun Chang, and Wen Hua (Eds.). Springer International Publishing, Cham, 642–658.

[33] Haoran Liu, Ying Zhang, Yujun Li, and Xiangyi Kong. 2021. Review on Emotion Recognition Based on Electroencephalography. *Frontiers in Computational Neuroscience* 15, October (2021), 1–15. https://doi.org/10.3389/fncom.2021.758212

[34] Therese Hoff Macan. 1994. Time Management: Test of a Process Model. *Journal of Applied Psychology* 79 (1994), 381–391. Issue 3. https://doi.org/10.1037/0021-9010.79.3.381

[35] Ryosuke Mita, Toshiki Taleuchi, Tomohiro Tanikawa, Takuji Narumi, and Michitaka Hirose. 2017. Early warning of task failure using task processing logs. *ACM International Conference Proceeding Series.* https://doi.org/10.1145/3041164.3041180

[36] Asuka Miyake, Masami Takahashi, Ryo Hashimoto, and Momoko Nakatani. 2021. StepUp Forecast: Predicting Future to Promote Walking. *Proceedings of MobileHCI 2021 - ACM International Conference on Mobile Human-Computer Interaction: Mobile Apart, MobileTogether* (2021). https://doi.org/10.1145/3447526.3472020

[37] Sean A. Munson, Hasan Cavusoglu, Larry Frisch, and Sidney Fels. 2013. Sociotechnical challenges and progress in using social media for health. *Journal of Medical Internet Research* 15, 10 (2013), 1–14. https://doi.org/10.2196/jmir.2792

[38] Jari-Erik Nurmi. 2005. *Thinking About and Acting Upon the Future: Development of Future Orientation Across the Life Span.* Lawrence Erlbaum Associates Publishers. 31–57 pages.

[39] Achim Peters, Bruce S. McEwen, and Karl Friston. 2017. Uncertainty and stress: Why it causes diseases and how it is mastered by the brain. *Progress in Neurobiology* 156 (2017), 164–188. https://doi.org/10.1016/j.pneurobio.2017.05.004

[40] Saeyoung Rho, Injung Lee, Hankyung Kim, Jonghyuk Jung, Hyungi Kim, Bong Gwan Jun, and Youn Kyung Lim. 2017. FutureSelf: What happens when we forecast self-trackers' future health statuses? *DIS 2017 - Proceedings of the 2017 ACM Conference on Designing Interactive Systems* (2017), 637–648. https://doi.org/10.1145/3064663.3064676

[41] A. John Rush, Madhukar H. Trivedi, Hicham M. Ibrahim, Thomas J. Carmody, Bruce Arnow, Daniel N. Klein, John C. Markowitz, Philip T. Ninan, Susan Kornstein, Rachel Manber, Michael E. Thase, James H. Kocsis, and Martin B. Keller. 2003. The 16-item Quick Inventory of Depressive Symptomatology (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR): A psychometric evaluation in patients with chronic major depression. *Biological Psychiatry* 54, 5 (sep 2003), 573–583. https://doi.org/10.1016/S0006-3223(02)01866-8

[42] Akane Sano, Amy Z Yu, Andrew W McHill, Andrew J K Phillips, Sara Taylor, Natasha Jaques, Elizabeth B Klerman, and Rosalind W Picard. 2015. Prediction of Happy-Sad mood from daily behaviors and previous sleep history. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS* 2015-Novem (2015), 6796–6799. https://doi.org/10.1109/EMBC.2015.7319954

[43] Sandra Servia-Rodríguez, Kiran K. Rachuri, Cecilia Mascolo, Peter J. Rentfrow, Neal Lathia, and Gillian M. Sandstrom. 2017. Mobile sensing at the service of mental well-being: A large-scale longitudinal study. In *26th International World Wide Web Conference, WWW 2017 (WWW '17)*. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, CHE, 103–112. https://doi.org/10.1145/3038912.3052618

[44] Lin Shu, Yang Yu, Wenzhuo Chen, Haoqiang Hua, Qin Li, Jianxiu Jin, and Xiangmin Xu. 2020. Wearable emotion recognition using heart rate data from a smart bracelet. *Sensors (Switzerland)* 20, 3 (2020), 1–19. https://doi.org/10.3390/s20030718

[45] Dimitris Spathis, Sandra Servia-Rodriguez, Katayoun Farrahi, Cecilia Mascolo, and Jason Rentfrow. 2019. Sequence multi-task learning to forecast mental wellbeing from sparse self-reported data. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (2019), 2886–2894. https://doi.org/10.1145/3292500.3330730

[46] Aaron Springer, Victoria Hollis, and Steve Whittaker. 2017. Dice in the Black Box: User Experiences with an Inscrutable Algorithm. www.aaai.org

[47] Toshiki Takeuchi, Takuji Narumi, Kyohei Suwa, Tomohiro Tanikawa, Hiroto Tamura, and Michitaka Hirose. 2013. A task-management system using future prediction based on personal lifelogs and plans. *UbiComp 2013 Adjunct - Adjunct Publication of the 2013 ACM Conference on Ubiquitous Computing*, 235–238. https://doi.org/10.1145/2494091.2494166

[48] Naoki Tateyama, Rui Fukui, and Shin'ichi Warisawa. 2022. Mood prediction based on calendar events using multitask learning. *IEEE Access* 10 (2022), –. https://doi.org/10.1109/ACCESS.2022.3193778

[49] Sara Taylor, Natasha Jaques, Ehimwenma Nosakhare, Akane Sano, and Rosalind Picard. 2020. Personalized Multitask Learning for Predicting Tomorrow's Mood, Stress, and Health. *IEEE Transactions on Affective Computing* 11, 2 (2020), 200–213. https://doi.org/10.1109/TAFFC.2017.2784832

[50] Anne E. Van Giezen, Ella Arensman, Philip Spinhoven, and Gezinus Wolters. 2005. Consistency of memory for emotionally arousing events: A review of prospective and experimental studies. *Clinical Psychology Review* 25, 7 (2005), 935–953. https://doi.org/10.1016/j.cpr.2005.04.011

[51] Jeffrey Warshaw, Tara Matthews, Steve Whittaker, Chris Kau, Mateo Bengualid, and Barton A. Smith. 2015. Can an algorithm know the "real you"? Understanding people's reactions to hyper-personal analytics systems. *Conference on Human Factors in Computing Systems - Proceedings* 2015-April, 797–806. https://doi.org/10.1145/2702123.2702274

[52] Timothy D. Wilson and Daniel T. Gilbert. 2003. Affective Forecasting. *Advances in Experimental Social Psychology* 35 (2003), 345–411. https://doi.org/10.1016/S0065-2601(03)01006-2

[53] Jianqiang Xu, Zhujiao Hu, Junzhong Zou, and Anqi Bi. 2020. Intelligent Emotion Detection Method Based on Deep Learning in Medical and Health Data. *IEEE Access* 8 (2020), 3802–3811. https://doi.org/10.1109/ACCESS.2019.2961139