Personality in Daily Life: Multi-Situational Physiological Signals Reflect Big-Five Personality Traits

Xinyu Shui, Yiling Chen, Xin Hu, Fei Wang, and Dan Zhang, *Member, IEEE*

***Abstract*—The popularity of wearable physiological recording devices has opened up new possibilities for the assessment of personality traits in everyday life. Compared with traditional questionnaires or laboratory assessments, wearable device-based measurements can collect rich data about individual physiological activities in real-life situations without interfering with normal life, enabling a more comprehensive description of individual differences. The present study aimed to explore the assessment of individuals' Big-Five personality traits by physiological signals in daily life situations. A commercial bracelet was used to track the heart rate (HR) data from eighty college students (all male) enrolled in a special training program with a strictly-controlled daily schedule for ten consecutive working days. Their HR activities were divided into five daily situations (morning exercise, morning classes, afternoon classes, free time in the evening, and self-study situations) according to their daily schedule. Regression analyses with HR-based features in these five situations averaged across the ten days revealed significant cross-validated quantitative prediction correlations of 0.32 and 0.26 for the dimensions of Openness and Extraversion, with the prediction correlation trending significance for Conscientiousness and Neuroticism. Moreover, the multi-situation HR-based results were in general superior to those based on single-situation HR-based features, as well as those based on the multi-situation self-reported emotion ratings. Togetherour findings demonstrate the link between personality and daily HR measures using state-of-the-art commercial devices and could shed light on the development of Big-Five personality assessment based on daily multi-situation physiological measures.**

***Index Terms*—Daily life, heart rate, multi-situation, personality, wearable device.**

Manuscript received 8 August 2022; revised 5 December 2022 and 25 January 2023; accepted 28 February 2023. Date of publication 8 March 2023; date of current version 6 June 2023. This work was supported in part by the National Natural Science Foundation of China under Grant 61977041, and in part by the Tsinghua University Spring Breeze Fund under Grant 2021Z99CFY037. *(Corresponding author: Dan Zhang.)*

Xinyu Shui, Yiling Chen, and Xin Hu are with the Department of Psychology, School of Social Sciences, Tsinghua University, Beijing 100084, China (e-mail: 602051386@qq.com; yiling-c19@tsinghua.org.cn; huxin530@gmail.com).

Fei Wang and Dan Zhang are with the Department of Psychology, School of Social Sciences, and Tsinghua Laboratory of Brain and Intelligence, Tsinghua University, Beijing 100084, China (e-mail: wf3126@mail.tsinghua.edu.cn; dzhang@tsinghua.edu.cn).

Digital Object Identifier 10.1109/JBHI.2023.3253820

I. Introduction

PERSONALITY refers to individual differences in dispositional patterns of thinking, feeling, and behaving that are related to important life outcomes such as career success, mental health, etc. [1], [2], [3]. The most commonly used personality trait framework is the Five-Factor Model which describes people in five dimensions, including Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism, [4], [5], [6]. While self-assessment questionnaires have been and still are the predominant way to conduct personality measurement, automatic personality assessment is gaining increasing attention in recent years [7], [8]. The working principle underlying such an approach is that one's dispositional personality traits are represented in the person's daily life. Data on daily life representations, e.g., behavior, corresponding to dispositional personality traits can, in turn, be assessed using the latest information technology such as social media, mobile phones, video cameras, wristbands, etc. [9], [10]. Compared to the classical questionnaire-based methods, the automatic approach is less likely to suffer from issues like social desirability, and self-presentational need [11], [12], especially in situations when their own personal interest is involved (e.g., job application).

The increasing popularity of the physiological recording technique has given rise to the physiological computing of personality [13], [14]. Compared to data sources such as social media and mobile phones, physiological data can continuously represent information related to an individual's activities in an objective way, without the requirement of one's active participation [15]. Among the commonly used physiological signals, the cardiovascular signal is one of the most effective signals in characterizing personality [16]. Using electrocardiogram (ECG) signals collected in the laboratory, researchers have demonstrated the association between individuals' cardiovascular activities and the Big-Five personality traits [17]. Higher Openness, which refers to one's tolerance of new ideas, has been correlated with several HRV features such as lower SDNN Index, ULF power, RMS-SD, etc. [18]; Higher Conscientiousness, which refers to one's tendency to be self-disciplined, has been linked to lower inter-beat interval (IBI) in pay-for-performance tasks [19]; Higher Extraversion, which describes one's tendency to be social, active, and outgoing, has been shown to be related to higher heart rate variability (HRV) and lower cardiovascular reactivity, especially in stressful situations [20], [21]; Higher

Agreeableness, which describes one's tendency of being kind, considerate, likable, helpful, and forgiving, has been found associated with increased resting-state ECG amplitude ratio and heart rate reactivity during psychological stress testing [22], [23]; Higher Neuroticism, which represents one's tendency to be anxious, self-conscious and paranoid, has been associated with longer resting-state QT interval duration, lower resting-state ECG amplitude, and lower task-state HR responses to acute stress [22], [24], [25]. As the ECG signals are believed to reflect both the parasympathetic and sympathetic activities [26], the above findings could be explained by the individualized autonomic nerve system characteristics in the resting and task states. In sum, laboratory-based studies have demonstrated a strong link between cardiovascular-related physiological activities and dispositional personality traits.

More importantly, the rapid development of wearable devices (e.g., smart watches, wristbands) has made it possible to record cardiovascular signals beyond the laboratory, towards daily life situations [27], [28]. Given the inconvenience of portable recording of ECG signals, the available commercial wearable devices have mainly employed photoplethysmography (PPG) to measure one's heart rate. While personality computing based on commercial PPG signals has not been fully explored [29], [30], studies have shown that PPG in daily life situations could effectively reflect mental stress [31], panic disorder [32], and mental health [33]. Given that PPGs have been suggested to represent cardiovascular activities in a way similar to ECGs (although with reduced performances) [34], [35], it is promising to expect effective computing of one's personality traits using daily-life PPGs.

However, state-of-the-art physiological studies have not fully addressed the situational issue of describing one's dispositional trait. Although the person-situation debate has often emerged in past individual difference studies, recent research perspectives argued that situational factors can moderate the influence of personality traits on one's state-level behavioral or physiological outcomes [36], [37]. For instance, staying with our friends rather than strangers always elicits talkativeness and affection. While situationally-oriented researches demonstrate that situation could have substantial impacts on one's thinking [38], feeling [39], and behaving [40], [41], most studies (especially these laboratory-based studies) to date have employed a single fixed situation, e.g., watching videos, reading words, playing games, etc. [42], [43], [44]. The limited situational richness may make the portrayal of individual traits inadequate. In line with this view, multi-task [45], multi-condition [46], or multi-scenario [47] methods are becoming increasingly valued in laboratory-based studies in recent years, demonstrating an increase in the accuracy and robustness of established models. The results of these studies show that the influence of personality is multi-dimensional and the consideration of situation information is important and necessary to improve the effectiveness of personality assessment. These findings suggest that situational factors are receiving increasing attention in personality research and are expected to provide vital support for improving assessment effectiveness.

Daily life is a more colorful source of situations than the laboratory [48]. Therefore, moving toward daily life scenarios, which contain rich situational information [40], becomes especially important and gains increasing attention [49], [50], [51]. This is partly because everyday life contains rich stimuli categories that experiments in laboratory settings can hardly reconstruct. For example, not all emotional feelings can be effectively evoked in laboratory settings such as pride, shame, etc. Another possible reason is that the stimuli intensity in laboratory settings differs from real-life situations. For example, the heart rate response when watching a soccer game at home is much higher than in laboratory stress-inducing tasks, suggesting that real-life stress may be much more intense than observed in the laboratory [48]. There are studies showing that ambulatory recordings of behavioral and physiological data in daily life settings using mobile phones or wearable devices, could be used to distinguish one's dispositional traits such as adaptive and maladaptive traits, loneliness and social isolation, distress and affective dysregulation, etc. [49], [50], [51]. Recordings in daily-life situations are believed to capture richer and more realistic information about an individual's physiological responses than in laboratory settings [52], hereby a more accurate and effective description of one's dispositional traits could be expected.

Nevertheless, there are two major challenges to be overcome before applying the wearable physiological recording technique for personality assessment in daily life situations. First, while commercial wearable devices are preferred for daily-life long-term recordings for their popularity, cost-effectiveness, and application potentials such as less interference on people's daily life, they are inferior to research-grade devices in terms of measurable data modalities, technical parameters, and other metrics. The employed device for the present study, for example, could only measure and output one heart rate value per second, making the ECG algorithms commonly used in laboratory studies unusable. Second, because of the variability of daily-life situations, it is not easy to identify and categorize different situations. We can classify situations by obtaining more information about the environment [53], [54], but this would require substantial efforts on environment sensing techniques, customized machine learning algorithms, etc., which might not be easily achieved. Alternatively, we can conduct research by focusing on specific population groups with a regular rhythm of life and work, e.g., students, workers, etc. While these population groups have been used in previous studies for their organized situation structures [47], [48], the situation information has not been sufficiently exploited. For instance, researchers have employed general information such as the time of the day (e.g., morning, afternoon, evening, etc.) to have a relatively coarse situation definition but the participants' daily life activities were quite diverse [50]. Physiological data recordings for people with well-controlled and clear situational structures are expected to facilitate the understanding of the expression of personality in daily life and promote wearable-device-based personality assessment applications.

In the present study, we aim to establish an automatic personality assessment model by using physiological signals recorded in daily-life situations. Eighty male college students in a special training program participated in our experiment for a two-week time period. The training program with a strictly-controlled schedule offered a unique opportunity to conduct a multi-situation study for real-life physiological measurement and HR-based personality assessment. A commercial bracelet was used to track the heart rate (HR) data during their daily college life. The recorded data were organized into five major situations from morning to evening: morning exercises, morning classes, afternoon classes, free time in the evening, and evening self-study sessions. The HR data of these five situations were used to assess their Big-Five personality traits (Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism, denoted as O, C, E, A, N), which are the most widely-used personality dimensions [4], [5], [6]. Regression analyses were conducted to evaluate the possible contributions of the multi-situational HR data for Big-Five personality assessment. The HR-based models were also compared to models based on situational emotional self-assessment. Our results demonstrate the feasibility of Big-Five personality assessment based on daily heart rate measures.

II. Methods

A. Participants

Eighty college students (all male, ranging from 17 to 22 years old, mean age = 19.1) were recruited for the present study. All these students were from an engineering major undergraduate program, from their first to the third year of college (27, 29, and 24 from the first, second, and third year respectively). The study was approved by the local Ethics Committee of Tsinghua University (IRB approval reference: 201906), and written informed consents were obtained from all participants prior to the start of the study. In particular, the participants were explicitly informed about the anonymization process of their data and were instructed to provide their report as accurately as possible.

During the two-week recording time, the participants were enrolled in a special training program with a strictly-controlled daily schedule. Their daily life was organized into five major situations: The Morning Exercise (ME) situation was from 06:00 to 08:00 when the participants completed more than one hour of intense physical activity (i.e., running); the Morning Classes (MC) situation and the Afternoon Classes (AC) situation was from 08:00 to 12:00 and 13:30 to 17:00, when the participants went to their classes (different grades in different classrooms); the Free time (FT) in the evening was from 17:00 to 19:00, when the participants had dinner and enjoyed their free time; the Self-Study (SS) situation was from 19:00 to 21:00, when the participants attended self-studying classes for 2 hours (different grades in different classrooms). The remaining periods such as lunch break were not taken into consideration due to their relatively short length. None of the students skipped any of these situation-specific activities during the two-week period, confirmed by the teacher responsible for this training program. Please note that the division into these five major situations was empirical. Nevertheless, such a regular schedule provided a unique opportunity for a situation-specific analysis by having a clear and easy-to-define situation structure.

B. Daily Physiological Recordings

Each participant was assigned a bracelet to measure their heart rate. The bracelet was from a local manufacturer (Fizzo Inc., China) and it cost less than 400 RMB for each band (about 60 US Dollars). The relatively low price made it suitable for large-scale daily studies.

The battery life of the bracelet can support a whole-day continuous recording, with the output heart rate values at a 1 Hz sampling rate. Importantly, the reliability of the performance of the bracelet has been previously validated [55]: the heart rates measured by the Fizzo bracelet can maintain high consistency with those of the research-level devices (Polar Team 2 Pro) in both indoor (ICC = 0.99, MAE = 0.69 bpm) and outdoor (ICC = 0.75, MAE = −2.6 bpm) exercise scenarios comparing with general heart rate tracking algorithms [56].

C. Procedure

Prior to the physiological recordings, all the participants completed the 44-item Big Five Personality Inventory [57], which was delivered through the online questionnaire platform (WJX.cn) on their mobile phones.

During the two-week recording time, the participants were instructed to keep the bracelet worn from 06:00 in the morning to 21:00 in the evening from Monday to Friday, with only exceptions such as taking showers, swimming, etc. The bracelets were charged and the data were exported every night.

From 21:00 to 23:00 every weekday, a revised day reconstruction method (DRM) was used to collect the participants' emotional states during these five major situations [58], [59], [60]. Specifically, the participants were asked to sequentially recall their daily activities during these five major situations from morning to evening of that day and report the corresponding emotional feelings for each of the five situations, as required by the DRM procedure. A 10-item Positive and Negative Affect Schedule (PANAS) [61] was used to record emotion states on upset, tired, ashamed, nervous, afraid, interested, inspired, determined, attentive, and active, all with 5-point Likert Scales. The procedure is illustrated in Fig. 1. The anonymized data are available at: https://dx.doi.org/10.21227/03kt-jz81 (IEEE DataPort) [62].

D. Data Processing

The HR data were first reviewed for artifacts. The time points with HR values below 40 BPM were empirically marked as invalid. Afterward, the retained HR data were first standardized by subtracting one's average HR within one recording day. The purpose of this operation is to remove the possible influence of individual differences in baseline heart rate. The choice of subtracting (rather than dividing, z-score, etc.) is expected to preserve the absolute HR value of the difference, which is expected to be physiologically meaningful. These standardized HR data were then segmented according to the defined five time periods for the five major situations, resulting in five data segments per participant per day. Data segments with less than

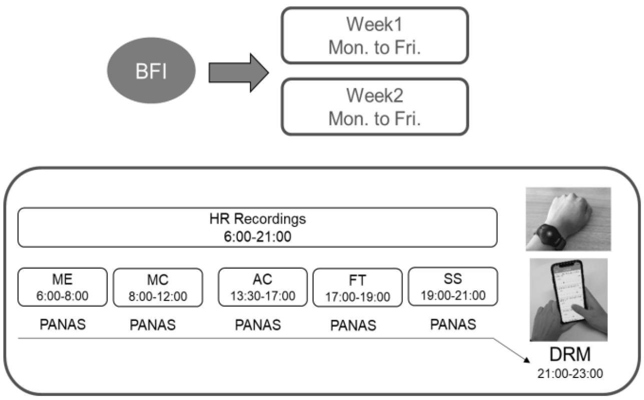

Fig. 1. Procedure of the two-week daily recording experiment. The participants finished Big-Five Inventory (BFI) before the two-week experiment. During the experiment, participants wore bracelet for Heart Rate (HR) recording and report their emotions through Positive and Negative Affect Schedule (PANAS) for five main situations (Morning Exercise (ME), Morning Classes (MC), Afternoon Classes (AC), Free time (FT) in the evening, and Self-Study (SS)) in a Day Reconstruction Method (DRM).

30 minutes of continuous valid data were excluded from further analysis, leaving relatively reliable and long-duration segments.

Based on previous literature [63], [64], [65], [66], the average level and variability of cardiovascular activity were frequently extracted as key features for psychological assessment, calculated as either the median and the interquartile range (denoted as HR-Median and HR-IQR), or the mean and the standard deviation (denoted as HR-Mean and HR-SD). Compared to HR-Mean and HR-SD, the HR-Median and HR-IQR are less sensitive to extreme values and therefore were expected to better reflect the physiological activities in different situations. Indeed, HR-Median and HR-IQR have been shown to be sensitive to situational factors such as psychological state, peer network, and social environment [67], [68]. In the present study, we calculated the median and the interquartile range of the HR data per segment as two features.

Further, the skewness and kurtosis of the HR data per segment were calculated and used as another two features as well (denoted as HR-Skew and HR-Kurt in the following) [69]. These features reflect the distribution of cardiovascular activity intensity over hours and, as higher-order spectra that indirectly reflect the physical state and emotional responses of the body, have been adapted in personalized emotion recognition and disease diagnosis [70], [71], [72]. These statistical indexes were calculated by using the data from all the available data segments.

The DRM-based self-reported emotion rating data were calculated for the scores of positive affect (PA) and negative affect (NA). Specifically, the PA score was defined as the mean of the five positive emotion scores (interested, inspired, determined, attentive, and active), and the NA score was defined as the mean of the five negative emotion scores (upset, tired, ashamed, nervous, and afraid). Using PA and NA scores instead of the ten items is a common practice [73], [74], as 1) PA and NA are emotional concepts with a clearer psychological meaning than these individual emotional items; 2) the PA and NA scores obtained by averaging multiple items had better measurement reliability than by these individual items. These HR features and the emotion rating features (PA and NA scores) were used for the following regression analyses.

E. Situation Specificity of the HR Features

To explore the situation specificity of the HR activities, one-way repeated-measures analysis of variance (ANOVA) with the factor situation (five major situations) on the four HR features, respectively. A significant effect on a specific HR feature would support its situation specificity. The one-way ANOVA was conducted for the emotion rating features as well and the comparison of the effect sizes would help us better understand the HR features.

F. Regression Analysis

To have a quantitative assessment of the personality score [9], [43], linear regression analyses were used to predict individual personality scores.

The main regression models (termed multi-situation models) were trained using the HR features (i.e., HR-Median, HR-IQR, HR-Skew, and HR-Kurt) from the five major situations, with a total number of twenty features (4 HR features $\times$ 5 situations = 20) per participant. By concatenating HR features from different situations, the combined feature set is expected to give a complete overview of the HR activities in all these situations. The HR features of each situation were further averaged across all 10 recording days to have a better representation of the situation-specific HR activities (see Results). A leave-one-participant-out cross-validation procedure was used to evaluate the performance of our regression models. The prediction model was constructed for each participant after leaving his or her own data out. Prediction accuracy was defined as the Pearson correlation coefficient between the predicted and self-report personality scores over all participants, as in previous personality computing studies [43].

Moreover, we trained multi-situation regression models with HR features extracted for each situation and averaged across the first N recording days ($N = 1, 2, \ldots, 9$), to explore the necessity of the practice of cross-day averaging (4 HR features $\times$ 5 situations = 20 features per participant, 9 models in total for averaging 1 to 9 days for each personality dimension).

To further explore the effectiveness of the main multi-situation regression models, the following regression models based on different features were trained for comparison:

1) Single-situation regression models with HR features for every single situation, to explore the necessity of having the multi-situation construct (4 HR features per participant, averaged across all the 10 recording days, 5 models in total for the 5 situations for each personality dimension).
2) All-situation regression models with HR features averaged across all situations, to explore the necessity of having a separate representation of the five situations rather than merging them together. Specifically, two averaging methods were considered for completeness. The first method calculated the HR features as the average of the ten-day averaged HR features in each situation

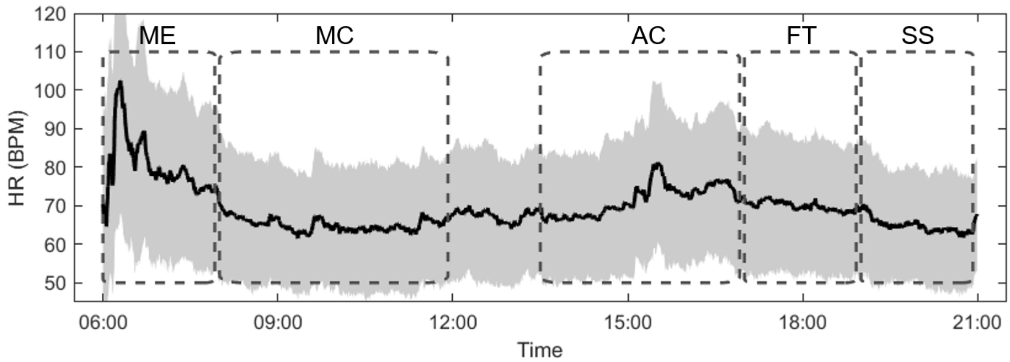

Fig. 2. The grand-average result of Heart Rate (HR) across all participants. The black line denotes the mean value, and the grey areas represent the standard deviation of the HR data at the participant level (the standard deviation of the average HR across all days within each participant).The dashed line showed five main situations: Morning Exercise (ME), Morning Classes (MC), Afternoon Classes (AC), Free time (FT) in the evening, and Self-Study (SS).

(4 HR features per participant, termed Method A). The second method calculated the HR features from raw HR data: HR-Median, HR-IQR, HR-Skew, and HR-Kurt were first extracted from HR data within each day across all situations, and then these features were defined as the average across all days (4 HR features $\times$ 1 situation per participant, termed Method B).

3) Multi-situation emotion rating regression models with the situation-specific emotion rating features averaged across all 10 days, to explore the possible contribution of the emotion rating features (2 emotion scores of PA and NA $\times$ 5 situations = 10 features per participant, one model for each personality dimension) for personality prediction.

To estimate the significance of these regression models' performance, a permutation test was employed. Specifically, similar regression models were trained with the above-defined features with randomly resampled labels, together with the same cross-validation procedure. The permutation method was repeated for 1000 times and the performances from these permutated data formed the null distribution for significance evaluation. Regression performance would be considered significant if the true performance was better than 95% of the permutated results. With the cross-validation procedure, only positive and correlation coefficients would be further evaluated for possible significance. A negative correlation coefficient indicates that the trained model yielded results in an opposite to expected direction on the testing data and hereby the model would not be valid.

Please note that although a different number of features were used in different regression models, the comparison is deemed sufficient as the cross-validated correlation was taken as the indicator of the regression performance.

III. Results

During the two-week recordings, the average recording time was 112.5 $\pm$ 35.0 hours (75.0 $\pm$ 16.8%, maximum 150 h) per participant, resulting in 9000.2 hours of HR data from all these 80 participants in 10 weekdays. For each situation, 516, 601, 666, 634, and 564 preprocessed HR segments were respectively obtained for ME, MC, AC, FT, and SS situation (maximum 800 segments). Meanwhile, a total number of 3, 805 PANAS reports were collected, with a mean response rate of 95.1%.

Fig. 2 shows the grand average result of HR data across all the recording days over all the participants. Their HRs increase rapidly from 06:00 when the participants started their morning exercises, reaching an average value of 100 BPM. Their HRs then steadily decrease to around 60-70 BPM throughout the day. The general trend of the HR data is consistent with intuitive expectations (especially the ME situation) and provides initial support for the validity of the collected data.

Table I shows the pairwise correlations between the HR features and the emotion rating features. The overall correlational results show significant correlations among the HR features (HR-Median, HR-IQR, HR-Skew, and HR-Kurt). The emotion ratings are not significantly correlated with the HR features, except for the weak correlation between PA and HR-Median ($r = -.11, p < .001$). Similar low emotion-HR correlations can be observed for the correlational results for each situation.

The one-way repeat-measures ANOVA reveals a significant effect of situations on all four HR features: HR-Median ($F(4, 316) = 260.3, p < .001, \eta^2 = .77$), HR-IQR ($F(4, 316) = 180.8, p < .001, \eta^2 = .70$), HR-Skew($F(4, 316) = 63.5, p < .001, \eta^2 = .45$), and HR-Kurt ($F(4, 316) = 86.1, p < .001, \eta^2 = .52$). The effect sizes of the HR features are comparable to those from the PA ($F(4, 316) = 68.2, p < .001, \eta^2 = .46$) and NA scores ($F(4, 316) = 6.5, p < .001, \eta^2 = .08$). These results suggest the situation specificity of the participants' emotion experiences and HR activities, providing support for the effectiveness of the five defined situations.

The pairwise correlational results between participants' personality scores and the corresponding features are summarized in Fig. 3. There were significant correlations between Extraversion and PA scores in FT and SS situations. A significant positive correlation was reported between Neuroticism and NA scores in all five situations, as well as significant negative correlations between Contentiousness and NA scores in each situation except Morning Exercise. Most of the correlation pairs between the HR features and personality remained insignificant except for five pairs: Openness and HR-Median (ME, FT, SS), Extraversion and HR-IQR (ME), and Extraversion and HR-Kurt (SS).

TABLE I
PAIRWISE CORRELATIONS BETWEEN THE HR FEATURES AND THE EMOTION RATINGS

$\bar{r}$	ALL	ME	MC	AC	FT	SS
HR-Median | HR-IQR	.53***	.39***	.46***	.54***	.50***	.46***
HR-Median | HR-Skew	−.55***	−.17***	−.24***	−.32***	−.37***	−.34***
HR-Median | HR-Kurt	−.42***	−.02	−.18***	−.30***	−.33***	−.29***
HR-IQR | HR-Skew	−.38***	.05	−.06	−.16***	−.18***	−.12***
HR-IQR | HR-Kurt	−.59***	−.11	−.24***	−.36***	−.38***	−.33***
HR-Skew | HR- Kurt	.62***	.28*	.47***	.54***	.57***	.57***
PA | NA	−.06	−.04	−.03	−.01	−.03	−.07
PA | HR-Median	−.11***	−.03	−.01	.04	.02	.01
PA | HR-IQR	.03	−.07	.00	.04	.03	.01
PA | HR-Skew	.03	−.08	−.11*	−.12***	−.09*	−.07*
PA | HR- Kurt	.04	.01	−.06	−.09*	−.05	−.04
NA | HR-Median	.04	−.02	.01	.03	.04	.02
NA | HR-IQR	.05	.01	.02	.04	.02	.02
NA | HR-Skew	−.02	−.02	.01	−.02	.00	−.01
NA | HR- Kurt	−.02	−.02	−.01	.03	−.03	−.02

Emotion Ratings: Positive Affects (PA), Negative Affects (NA); HR Features: Median, Interquartile-Range, Skewness and Kurtosis of Heart Rate (HR-Median, HR-IQR, HR-Skew, HR-Kurt).
* p < .05, *** p < .001, Bonferroni corrected. ALL: correlation over all situations; Five single situations: The Morning Exercise (ME), Morning Classes (MC), Afternoon Classes (AC), Free time (FT) in the evening, and Self-Study (SS).
The overall correlation coefficients were calculated first within each participant over all situations (i.e., N = 5 (situations) × 10 (days) = 50) and then averaged across the participants. A series of t-tests were employed to test whether the mean of the correlation coefficients across all the participants was significantly different from zero. The correlation coefficients were z-transformed before the averaging and t-tests.
The situation-specific correlation coefficients were calculated in a similar way, but only with the data from one single situation (i.e., N = 10 (days)).

Using the ten-day-averaged HR features, significant cross-validated regression accuracies (in the form of correlation coefficients) are observed for Openness ($r = .32$, $p < .01$, 95% CI: [.09, .54]), and Extraversion ($r = .26$, $p = .03$, 95% CI: [.03, .48]) in Fig. 4(a). The bootstrapped permutation method output 95% confidence interval of null distribution by repeating the regression algorithm 1000 times for shuffled prediction targets in Fig. 4(b). The regression accuracies on multiple personality dimensions show a clear increasing trend with the increasing number of recording days before feature averaging in Fig. 4(b). The prediction results of Openness and Extraversion reach a significant level ($p < .05$) with nine days or more data. A similar trend can be found in Neuroticism and Conscientiousness dimensions. It's worth noting that the prediction accuracy of Neuroticism ($r = .22$, $p = .06$) and Conscientiousness ($r = .12$, $p = .28$) are relatively high and gains large increase from five-day to ten-day data ($\Delta r_N = .37$; $\Delta r_C = .37$), although their final accuracies are not statistically significant.

Table II summarizes the regression results of the main multi-situation model, as well as the to-be-compared models, all based on ten-day-averaged features. While the prediction accuracies of multi-situation HR models are significant on Openness, and Extraversion dimensions, all the other models (including the all-situation models and single-situation models) yield insignificant results. Furthermore, Fig. 5 shows the comparison between the HR-based models and DRM-based models, both of which are based on multi-situational features. HR-based models raise a better performance on Openness (95% CI of ΔZ_r: [0.08, 0.73], significant), Conscientiousness (95% CI of ΔZ_r: [−0.16, 0.48]), and Extraversion (95% CI of ΔZ_r: [−0.01, 0.63]), while both models perform well in Neuroticism (DRM-based: $r = .21$, $p = .07$, 95% CI: [−.02, .44]).

IV. Summary and Discussion

In the present study, heart rate data were collected from a cohort of college students during their daily life during a special training program for a two-week time period. The special training program with a strictly-controlled schedule offered a unique opportunity to conduct a multi-situation study for heart

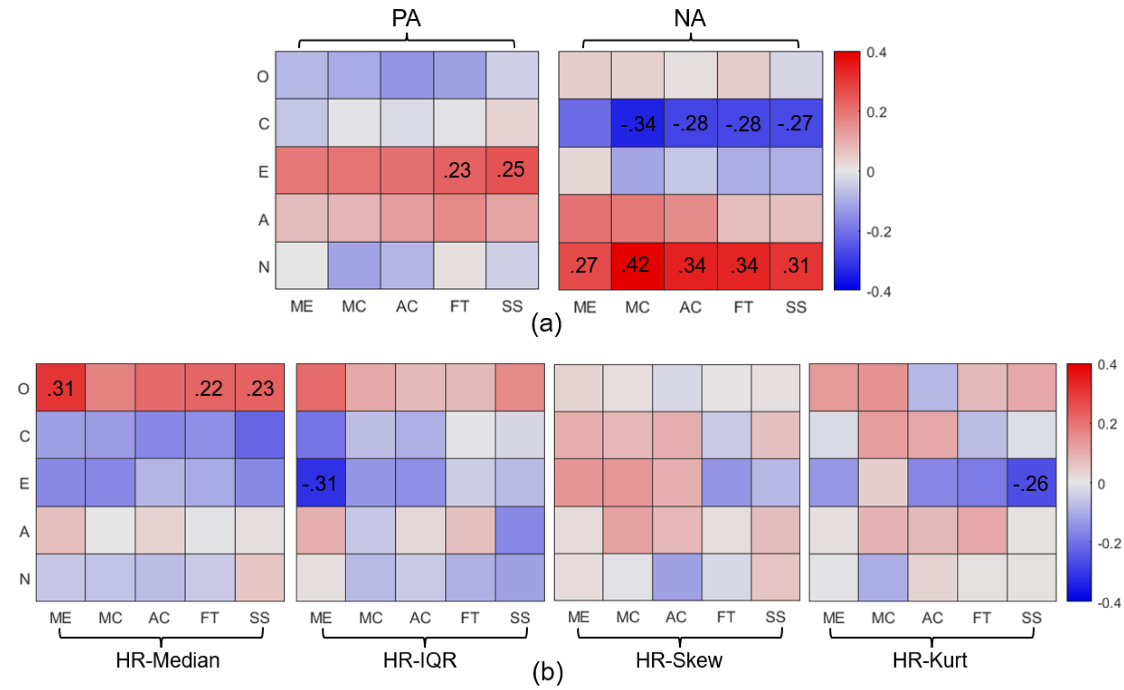

Fig. 3. (a) Pairwise correlation between Positive Affects (PA)/Negative Affects (NA) scores and Big-Five personality traits. O: Openness; C: Conscientiousness; E: Extraversion; A: Agreeableness; N: Neuroticism. Five single situations: Morning Exercise (ME), Morning Classes (MC), Afternoon Classes (AC), Free time (FT) in the evening, and Self-Study (SS). (b) Pairwise correlation between HR features and Big-Five personality traits. HR Features: Median, Interquartile-Range, Skewness and Kurtosis of Heart Rate (HR-Median, HR-IQR, HR-Skew, HR-Kurt). The PA/NA scores and HR features were first averaged across ten days before correlation analysis. Each pixel indicates the correlation pair between one averaged feature in each situation (80 $\times$ 1) and one dimension of Big-Five (80$\times$1). Correlation coefficients were marked on significant channels ($p < .05$). Only N-NA(MC) pair passed the Bonferroni test ($p < .001$, bold coefficient).

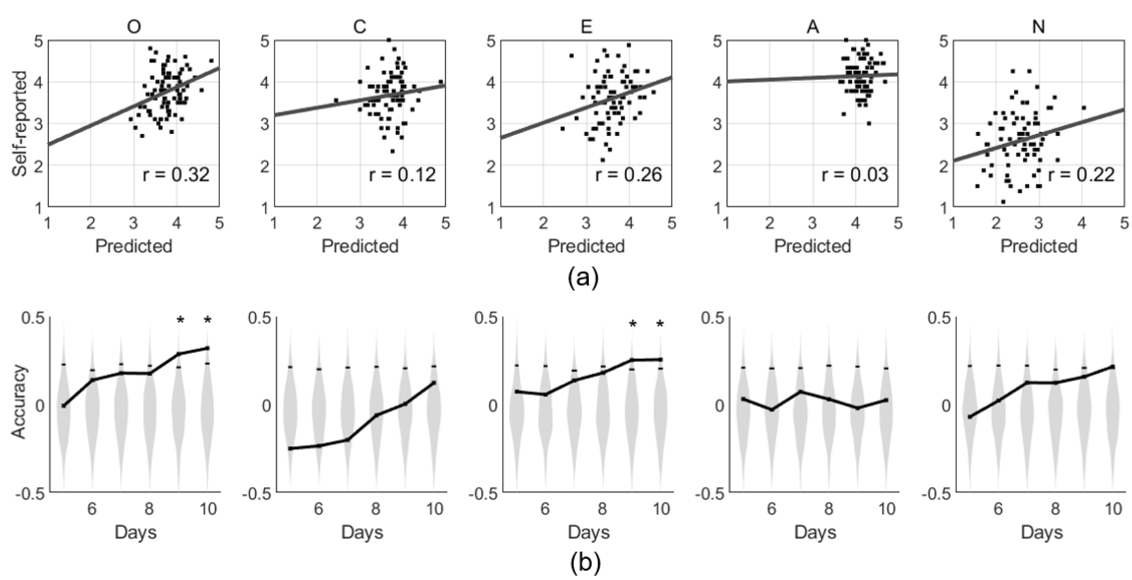

Fig. 4. (a) Scatterplots of the regression result of multi-situation models based on HR data. O: Openness; C: Conscientiousness; E: Extraversion; A: Agreeableness; N: Neuroticism. (b) The multi-situation regression results as the function of the number of days being averaged for the HR features. Violin plots show the distribution of accuracies obtained from the permutated regression models (null distribution). The short black lines in violin areas indicate the 95% confidence interval of null distributions. The asterisk (*) indicates a two-tailed significant regression accuracy at the .05 level.

TABLE II
COMPARISON BETWEEN MULTI-SITUATION AND SINGLE-SITUATION HR REGRESSION RESULT

Accuracy	O	C	E	A	N
MUL	**.32***	.12	**.26***	–.03	**.22**
ALL-A	.10	–.11	.04	–.29	–.24
ALL-B	–.03	–.17	–.01	–.14	–.34
ME	.19	–.14	.07	–.51	–.65
MC	.05	–.12	–.09	.03	–.29
AC	.20	–.05	.17	–.28	.01
FT	.03	–.18	–.11	–.22	–.45
SS	.05	.07	.13	–.01	–.04

O: Openness; C: Conscientiousness; E: Extraversion; A: Agreeableness; N: Neuroticism. The asterisk indicates significance.
MUL: the main multi-situation regression model; ALL-A and ALL-B: the all-situation models using methods A and B, see Methods; from ME to SS: the single-situation models using all the 5 situations.
Please note that the large negative correlation coefficients did not mean a large effect size: it only means no valid prediction in the model under the Leave-one-out framework.

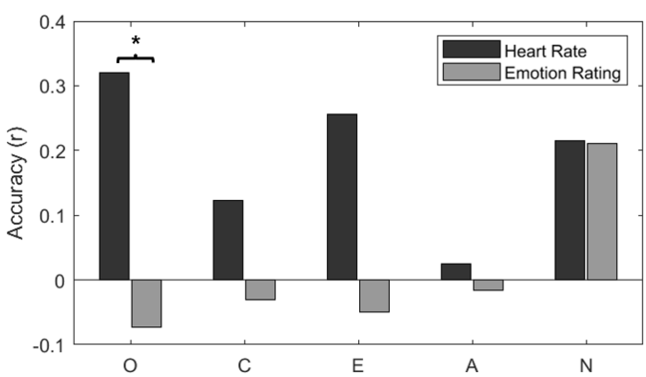

Fig. 5. The regression results based on the HR and Emotion rating features. O: Openness; C: Conscientiousness; E: Extraversion; A: Agreeableness; N: Neuroticism. Confidence intervals of correlation coefficients are calculated to test the difference in prediction accuracy between the two models. The asterisk (*) indicates a two-tailed significant difference at the .05 level. HR-based models raise significantly higher accuracy on Openness dimensions.

rate-based personality computing. By extracting HR features from five major daily-life situations and cross-validated regression method, the multi-situation HR models significantly predict personality scores for Extraversion and Openness. A clear increasing trend of the prediction accuracy as the function of the increasing number of recording days was observed for Openness and Extraversion. The multi-situation HR-based models showed better performance than single-situation HR-based models as well as multi-situation emotion rating-based models. Our findings demonstrate the feasibility of Big-Five personality assessment based on daily heart rate measures using state-of-the-art commercial devices and highlight the importance of the multi-situational construct.

While the regression results for the Big-Five personality dimensions of Openness and Extraversion are in general agreement with previous studies [22], [20], [68], our study further extends the state-of-the-art knowledge on how the situation factor could come into play. Rather than having a resting-state session or a single laboratory-based task (e.g., video watching, laboratory stress) [75], [76], we included heart rate data from multiple situations from daily life settings. While the non-significant model performance by these single-situation models seemed to contradict the reported moderating effect of the situation on the influence of personality traits on one's behavioral or physiological outcomes in literature [37], our results could be attributed to the relatively coarse measurement of the situation-specific physiological outcomes using the lightweight commercial device, or the relatively conservative but physiologically-plausible feature extraction and regression analysis methods. While further studies with advanced sensors and elaborated algorithms may help elucidate the situation-specific physiological activity patterns in daily scenarios, the superior performance of the multi-situation regression models as compared to the single-situation models demonstrates the necessity of having multiple situations from an application-oriented perspective: the combination of data from typical daily-life situations enabled a more complete description of the student, therefore supporting a more accurate personality assessment. In addition, our results demonstrate the feasibility of using lightweight commercial devices for daily recording and personality assessment by using a 400-RMB (approx. 60 US Dollars) commercial device that is representative of the performance of state-of-the-art lightweight devices. Given the increasing popularity of wearable devices [77], the positive results from such a typical device demonstrate the potential of personality assessment beyond the laboratory into daily environments.

It is important to note that the overall performance of regression models got improved with the increase of recording days being averaged for feature extraction. Given the limited physiological information (i.e., only heart rate) provided by the state-of-the-art commercial bracelet, it is reasonable to expect a longer time period needed for a reliable assessment, as compared to the laboratory-based research-level devices (e.g., high temporal resolution ECG devices). Nevertheless, we indeed observed significant situation specificity of the recorded HR features, with comparable effect sizes as the emotion ratings. Hereby, the within-situation consistency of the HR features could support more robust estimates of the situation-specific physiological activities and thus better regression performances [78], [79]. Considering the convenience of wearing such devices for a relatively long time period (e.g., for other life-related purposes given the rich functions integrated into wearable devices), it is not difficult to collect data for more than ten days for personality assessment. Therefore, better performances can be expected if more data are obtained. While the practice of feature averaging as adopted in the present study is a simple and straightforward way to enhance the signal-to-noise ratio of the situation-specific HR features, more advanced statistical or machine learning methods that could make better use of the nested data structure (i.e., HR features nested in situations, situations nested in days) are expected to further improve the performance. In addition, with the recording of data on a monthly or annual scale, we may potentially track the development of psychological traits for applications such as mental health monitoring [80].

It is also necessary to point out that the HR-based results outperformed DRM emotion ratings on several personality dimensions. Although the retrospective DRM-based emotion ratings could be potentially biased to some extent [81], [82], [83], the observed bivariate correlations between the emotion ratings and personality dimensions support the validity of the ratings in the present study. For instance, significant correlations between NA and Neuroticism were found in all five situations (Fig. 3(a)), as frequently reported in previous studies [84]. Interestingly, while a larger number of significant bivariate correlations were obtained for the PA/NA-personality relations as compared to the HR-personality relations, the regression analysis showed better overall results for the HR-based regression models than the PA/NA-based regression models. This piece of results would suggest complementarity among the multi-situational HR features, again arguing for the necessity of having the multi-situational construct. Moreover, the difference in the regression performances using the HR features and the emotion ratings would suggest richer information in the HR features for personality prediction beyond emotional experiences. On one hand, both features could reflect emotion-related activities but from different perspectives at least in some cases, e.g., for the prediction of Openness and Extraversion, which have been known to be closely related to emotion [85], [86] and evidenced in the present study (Fig. 3(a)). On the other hand, HR features might contain information related to attention during classroom learning, social anxiety, etc. [87], [88], which could potentially support the explanation of the Openness results that were not related to PA and NA in the present study.

The present study has some limitations that should be noted. First, the participants' daily schedule was empirically divided into five major situations according to the schedule of their special training program. The current situation definition was relatively coarse, and it could be further improved to obtain a more detailed description of the individuals, e.g., by introducing more sensors for more information about the individuals and their environment [89]. Second, self-report Big-Five Personality scores were used as the gold standard in the present study, as in most personality computing studies [43], [90]. Admittedly, the findings could be influenced by the self-report bias issue. The influence could be limited, as the participants in the present study were explicitly informed about the anonymization process of their data and hereby more accurate self-evaluation would be expected. Third, it needs to be noted that the participants are all male due to the special training program. As gender differences in personality have been widely acknowledged [91], [92], the present findings could be limited in extending to females. Therefore, it would be necessary to conduct further studies with female participants in order to obtain a complete overview of how personality is linked to daily physiological measures. Last but not least, as possibly the first personality computing study using daily-life multi-situation data from a commercially available bracelet, caution must be taken when interpreting the results [93], [94]. Especially, the reliability of the results needs to be validated in future studies, ideally with advanced technologies to collect multi-situational physiological data (see below).

The findings of this study are expected to have clearer applications as technology advances. Since two out of the five personality dimensions (Openness, Extraversion) showed significant regression results (with trending significance in the dimension of Conscientiousness and Neuroticism), previous physiological studies in laboratory settings have suggested the possibility of assessing all five personality traits [29], [43], [95]. As the rapid growth of this field is making more and more signal sensing (e.g., easy-to-wear ECG [96], ear-based EEG, wrist-worn skin temperature, wrist-worn skin conductance [97], etc.) possible in daily-life situations [98], [99], [100], the performance of daily-life-based personality computing assessment is expected to be greatly promoted. Moreover, the multi-situation construct is expected to be extended to adapt to more scenarios beyond the special training program as in the present study. On one hand, the present methodological framework could be applied to people with relatively fixed schedules, e.g., factory workers, primary and middle school students, etc. On the other hand, for people with more flexible schedules, smart sensing of their situations could be achieved with the help of mobile phones, wearable and ambient sensors, as well as environment sensors embedded in the home and workplace, etc. [101], [102], [103], [104]. By defining proper situations (e.g., business meetings, shopping, doing housework, etc.) out of their daily life, multi-situation personality computing models based on these sensor data could be implemented. In sum, technological advances are expected to greatly promote the description of an individual's dispositional traits and personality assessments are expected to meet the needs of practicality in the near future. The present study contributes to this prospect by proposing the multi-situation personality computing framework and preliminarily demonstrating its feasibility using daily-life physiological sensing.

References

[1] D. J. Ozer and V. Benet-Martínez, "Personality and the prediction of consequential outcomes," *Annu. Rev. Psychol.*, vol. 57, pp. 401–421, 2006.

[2] J. Lodi-Smith and B. W. Roberts, "Social investment and personality: A meta-analysis of the relationship of personality traits to investment in work, family, religion, and volunteerism," *Pers. Social Psychol. Rev.*, vol. 11, no. 1, pp. 68–86, 2007.

[3] P. J. Corr and G. Matthews, Eds., *The Cambridge Handbook of Personality Psychology*. Cambridge, U.K.: Cambridge Univ. Press, 2009.

[4] O. P. John, L. P. Naumann, and C. J. Soto, "Paradigm shift to the integrative big five trait taxonomy: History, measurement, and conceptual issues," in *Handbook of Personality: Theory and Research*, 3rd ed. New York, NY, USA: Guilford Press, 2008, pp. 114–158.

[5] P. T. Costa and R. R. McCrae, "Four ways five factors are basic," *Pers. Individual Differences*, vol. 13, no. 6, pp. 653–665, 1992.

[6] R. R. McCrae and P. T. Costa Jr., "Empirical and theoretical status of the five-factor model of personality traits," in *The SAGE Handbook of Personality Theory and Assessment, Vol 1: Personality Theories and Models*. Thousand Oaks, CA, USA: Sage Publications, 2008, pp. 273–294.

[7] S. Vazire, "Informant reports: A cheap, fast, and easy method for personality assessment," *J. Res. Pers.*, vol. 40, no. 5, pp. 472–481, 2006.

[8] S. V. Kedar and D. S. Bormane, "Automatic personality assessment: A systematic review," in *Proc. IEEE Int. Conf. Inf. Process.*, 2015, pp. 326–331.

[9] W. Youyou, M. Kosinski, and D. Stillwell, "Computer-based personality judgments are more accurate than those made by humans," *Proc. Nat. Acad. Sci.*, vol. 112, no. 4, pp. 1036–1040, 2015.

[10] L. V. Phan and J. F. Rauthmann, "Personality computing: New frontiers in personality assessment," *Social Pers. Psychol. Compass*, vol. 15, no. 7, 2021, Art. no. e12624.

[11] M. L. Kern et al., "Social media-predicted personality traits and values can help match people to their ideal jobs," *Proc. Nat. Acad. Sci.*, vol. 116, pp. 26459–26464, 2019.

[12] C. Viswesvaran and D. S. Ones, "Meta-analyses of fakability estimates: Implications for personality measurement," *Educ. Psychol. Meas.*, vol. 59, no. 2, pp. 197–210, 1999.

[13] Z. Ihsan and A. Furnham, "The new technologies in personality assessment: A review," *Consulting Psychol. J.: Pract. Res.*, vol. 70, no. 2, pp. 147–166, 2018.

[14] A. Vinciarelli and G. Mohammadi, "A survey of personality computing," *IEEE Trans. Affect. Comput.*, vol. 5, no. 3, pp. 273–291, Jul.–Sep. 2014.

[15] V. Vijayan et al., "Review of wearable devices and data collection considerations for connected health," *Sensors*, vol. 21, no. 16, 2021, Art. no. 5589.

[16] S. Israel et al., "ECG to identify individuals," *Pattern Recognit.*, vol. 38, no. 1, pp. 133–142, 2005.

[17] T. Boljanić et al., "Relationship between electrocardiogram-based features and personality traits: Machine learning approach," *Ann. Noninvasive Electrocardiol.*, vol. 27, no. 1, 2022, Art. no. e12919.

[18] A. Zohar, C. Cloninger, and R. McCraty, "Personality and heart rate variability: Exploring pathways from personality to cardiac coherence and health," *Open J. Social Sci.*, vol. 1, no. 6, pp. 32–39, 2013.

[19] K. L. Harper et al., "Conscientiousness and effort-related cardiac activity in response to piece-rate cash incentives," *Motivation Emotion*, vol. 42, no. 3, pp. 377–385, 2018.

[20] C. R. Jonassaint et al., "The effects of neuroticism and extraversion on cardiovascular reactivity during a mental and an emotional stress task," *Int. J. Psychophysiol.*, vol. 74, no. 3, pp. 274–279, 2009.

[21] O. Soliemanifar, A. Soleymanifar, and R. Afrisham, "Relationship between personality and biological reactivity to stress: A review," *Psychiatry Investigation*, vol. 15, no. 12, pp. 1100–1114, 2018.

[22] S. Koelsch, J. Enge, and S. Jentschke, "Cardiac signatures of personality," *PLoS One*, vol. 7, no. 2, 2012, Art. no. e31441.

[23] A. Bibbey et al., "Personality and physiological reactions to acute psychological stress," *Int. J. Psychophysiol.*, vol. 90, no. 1, pp. 28–36, 2013.

[24] P. Minoretti et al., "QT interval duration in apparently healthy men is associated with depression-related personality trait neuroticism," *J. Psychosomatic Res.*, vol. 61, no. 1, pp. 19–23, 2006.

[25] Y. Xin et al., "The relationship between personality and the response to acute psychological stress," *Sci. Rep.*, vol. 7, pp. 1–8, 2017.

[26] L. Fedele and T. Brand, "The intrinsic cardiac nervous system and its role in cardiac pacemaking and conduction," *J. Cardiovasc. Develop. Dis.*, vol. 7, no. 4, 2020, Art. no. 54.

[27] H. C. Ates et al., "End-to-end design of wearable sensors," *Nature Rev. Mater.*, vol. 7, no. 11, pp. 887–907, 2022.

[28] A. J. Perez and S. Zeadally, "Recent advances in wearable sensing technologies," *Sensors*, vol. 21, no. 20, 2021, Art. no. 6828.

[29] I. Liu, S. Ni, and K. Peng, "Predicting personality with smartphone cameras: A Pilot study," in *Proc. IEEE Int. Conf. Hum.-Mach. Syst.*, 2020, pp. 1–6.

[30] A. R. Butt, A. Arsalan, and M. Majid, "Multimodal personality trait recognition using wearable sensors in response to public speaking," *IEEE Sensors J.*, vol. 20, no. 12, pp. 6532–6541, Jun. 2020.

[31] S. Gedam and S. Paul, "A review on mental stress detection using wearable sensors and machine learning techniques," *IEEE Access*, vol. 9, pp. 84045–84066, 2021.

[32] K. Na, S. Cho, and S. Cho, "Machine learning-based discrimination of panic disorder from other anxiety disorders," *J. Affect. Disord.*, vol. 278, pp. 1–4, 2021.

[33] H. W. Loh et al., "Application of photoplethysmography signals for healthcare systems: An in-depth review," *Comput. Methods Prog. Biomed.*, vol. 216, 2022, Art. no. 106677.

[34] M. Bolanos, H. Nazeran, and E. Haltiwanger, "Comparison of heart rate variability signal features derived from electrocardiography and photoplethysmography in healthy individuals," in *Proc. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2006, pp. 4289–4294.

[35] D. Castaneda et al., "A review on wearable photoplethysmography sensors and their potential future applications in health care," *Int. J. Biosensors Bioelectron.*, vol. 4, no. 4, pp. 195–202, 2018.

[36] R. M. Furr and D. C. Funder, "Persons, situations, and person–situation interactions," in *Handbook of Personality: Theory and Research*, O. P. John and R. W. Robins, Eds., 4th ed. New York, NY, USA: Guilford Press, 2021, pp. 568–580.

[37] N. Kuper, A. S. von Garrel, B. M. Wiernik, L. V. Phan, N. Modersitzki, and J. Rauthmann, "Distinguishing four types of person $\times$ situation interactions: An integrative framework and empirical examination." *PsyArXiv*, 2022, doi: 10.31234/OSF.IO/2G63C.

[38] C. S. Dweck, C. Chiu, and Y. Hong, "Implicit theories and their role in judgments and reactions: A world from two perspectives," *Psychol. Inquiry*, vol. 6, no. 4, pp. 267–285, 1995.

[39] J. Slaby, "Mind invasion: Situated affectivity and the corporate life hack," *Front. Psychol.*, vol. 7, 2016, Art. no. 266.

[40] E. Skimina and J. Cieciuch, "Explaining everyday behaviours and situational context by personality metatraits and higher-order values," *Eur. J. Pers.*, vol. 34, no. 1, pp. 29–59, 2020.

[41] R. A. Sherman et al., "The independent effects of personality and situations on real-time expressions of behavior and emotion," *J. Pers. Social Psychol.*, vol. 109, no. 5, pp. 872–888, 2015.

[42] X. Hu, F. Wang, and D. Zhang, "Similar brains blend emotion in similar ways: Neural representations of individual difference in emotion profiles," *Neuroimage*, vol. 247, 2022, Art. no. 118819.

[43] W. Li et al., "Quantitative personality predictions from a brief EEG recording," *IEEE Trans. Affect. Comput.*, vol. 13, no. 3, pp. 1514–1527, Jul.–Sep. 2022, doi: 10.1109/TAFFC.2020.3008775.

[44] W. Li et al., "EEG responses to emotional videos can quantitatively predict big-five personality traits," *Neurocomputing*, vol. 415, pp. 368–381, 2020.

[45] A. K. Ahmet, L. Akarun, and O. Aran, "Multi-domain and multi-task prediction of extraversion and leadership from meeting videos," *EURASIP J. Image Video Process.*, vol. 2017, no. 1, pp. 1–14, 2017.

[46] Y. Mehta et al., "Recent trends in deep learning based personality detection," *Artif. Intell. Rev.*, vol. 53, no. 4, pp. 2313–2339, 2020.

[47] D. Dotti, M. Popa, and S. Asteriadis, "Being the center of attention: A person-context CNN framework for personality recognition," *ACM Trans. Interactive Intell. Syst.*, vol. 10, no. 3, pp. 1–20, 2020.

[48] F. H. Wilhelm and P. Grossman, "Emotions beyond the laboratory: Theoretical fundaments, study design, and analytic strategies for advanced ambulatory assessment," *Biol. Psychol.*, vol. 84, no. 3, pp. 552–569, 2010.

[49] R. Yan et al., "Exploratory machine learning modeling of adaptive and maladaptive personality traits from passively sensed behavior," *Future Gener. Comput. Syst.*, vol. 132, pp. 266–281, 2022.

[50] A. Doryab et al., "Identifying behavioral phenotypes of loneliness and social isolation with passive sensing: Statistical analysis, data mining and machine learning of smartphone and fitbit data," *JMIR mHealth uHealth*, vol. 7, no. 7, Jul. 2019, Art. no. e13209.

[51] U. Ebner-Priemer et al., "Distress and affective dysregulation in patients with borderline personality disorder: A psychophysiological ambulatory monitoring study," *J. Nervous Ment. Dis.*, vol. 196, no. 4, pp. 314–320, 2008.

[52] D. Trampe, J. Quoidbach, and M. Taquet, "Emotions in everyday life," *PLoS One*, vol. 10, no. 12, 2015, Art. no. e0145450.

[53] V. Bijalwan, B. S. Vijay, and V. Gupta, "Wearable sensor-based pattern mining for human activity recognition: Deep learning approach," *Ind. Robot*, vol. 49, no. 1, pp. 21–33, 2022.

[54] T. Ivascu, K. Cincar, A. Dinis, and V. Negru, "Activities of daily living and falls recognition and classification from the wearable sensors data," in *Proc. IEEE E-Health Bioeng. Conf.*, 2017, pp. 627–630.

[55] J. Sun and Y. Liu, "Using smart bracelets to assess heart rate among students during physical education lessons: Feasibility, reliability, and validity study," *JMIR mHealth uHealth*, vol. 8, no. 8, 2020, Art. no. e17699.

[56] J. Guo, X. Chen, J. Zhao, X. Zhang, and X. Chen, "An effective photoplethysmography heart rate estimation framework integrating two-level denoising method and heart rate tracking algorithm guided by finite state machine," *IEEE J. Biomed. Health Inform.*, vol. 26, no. 8, pp. 3731–3742, Aug. 2022.

[57] R. Carciofo et al., "Psychometric evaluation of chinese language 44-item and 10-item big five personality inventories, including correlations with chronotype, mindfulness and mind wandering," *PLoS One*, vol. 11, no. 2, 2016, Art. no. .e0149963.

[58] D. Kahneman et al., “A survey method for characterizing daily life experience: The day reconstruction method,” *Science*, vol. 306, no. 5702, pp. 1776–1780, 2004.

[59] X. Shui et al., “A dataset of daily ambulatory psychological and physiological recording for emotion research,” *Sci. Data*, vol. 8, no. 1, 2021, Art. no. 161.

[60] K. Räikkönen et al., “Effects of optimism, pessimism, and trait anxiety on ambulatory blood pressure and mood during everyday life,” *J. Pers. Social Psychol.*, vol. 76, no. 1, pp. 104–113, 1999.

[61] D. Watson, L. A. Clark, and A. Tellegen, “Development and validation of brief measures of positive and negative affect: The PANAS scales,” *J. Pers. Social Psychol.*, vol. 54, no. 6, pp. 1063–1070, 1988.

[62] X. Shui, “Personality in daily life: Multi-situational physiological signals reflect big-five personality traits,” *IEEE DataPort*, doi: 10.21227/03kt-jz81.

[63] C. C. Brumbaugh et al., “Physiological correlates of the Big 5: Autonomic responses to video presentations,” *Appl. Psychophysiol. Biofeedback*, vol. 38, no. 4, pp. 293–301, 2013.

[64] S. Lin et al., “Wearable emotion recognition using heart rate data from a smart bracelet,” *Sensors*, vol. 20, no. 3, 2020, Art. no. 718.

[65] T. W. Kamarck et al., “Effects of task strain, social conflict, and emotional activation on ambulatory cardiovascular activity: Daily life consequences of recurring stress in a multiethnic adult sample,” *Health Psychol.*, vol. 17, no. 1, pp. 17–29, 1998.

[66] M. Schneider et al., “Life events are associated with elevated heart rate and reduced heart complexity to acute psychological stress,” *Biol. Psychol.*, vol. 163, 2021, Art. no. 108116.

[67] C. Wang, O. Lizardo, and D. S. Hachen, “Using fitbit data to monitor the heart rate evolution patterns of college students,” *J. Amer. College Health*, vol. 70, pp. 875–882, 2022.

[68] A.-M. Brouwer, M. G. van Schaik, J. E. Korteling, J. B. F. van Erp, and A. Toet, “Neuroticism, extraversion, conscientiousness and stress: Physiological correlates,” *IEEE Trans. Affect. Comput.*, vol. 6, no. 2, pp. 109–117, Apr.–Jun. 2015.

[69] K. C. Chua et al., “Application of higher order statistics/spectra in biomedical signals—A review,” *Med. Eng. Phys.*, vol. 32, no. 7, pp. 679–689, 2010.

[70] A. Shapiro et al., “A novel digital approach to describe real world outcomes among patients with constipation,” *NPJ Digit. Med.*, vol. 4, no. 1, 2021, Art. no. 27.

[71] I. A. Karaye, S. Saminu, and N. Ozkurt, “Analysis of cardiac beats using higher order spectra,” in *Proc. IEEE 6th Int. Conf. Adaptive Sci. Technol.*, 2014, pp. 1–8.

[72] S. Zhao et al., “Personalized emotion recognition by personality-aware high-order learning of physiological signals,” *ACM Trans. Multimedia Comput., Commun., Appl.*, vol. 15, no. 1, pp. 1–18, 2019.

[73] A. Steptoe and J. Wardle, “Positive affect measured using ecological momentary assessment and survival in older men and women,” *Proc. Nat. Acad. Sci.*, vol. 108, no. 45, 2011, Art. no. 18244.

[74] E. L. Merz and S. C. Roesch, “Modeling trait and state variation using multilevel factor analysis with PANAS daily diary data,” *J. Res. Pers.*, vol. 45, no. 1, pp. 2–9, 2011.

[75] D. C. Schwebel and J. Suls, “Cardiovascular reactivity and neurotiscm: Results from a laboratory and controlled ambulatory stress protocol,” *J. Pers.*, vol. 67, no. 1, pp. 67–92, 1999.

[76] R. Subramanian, J. Wache, M. K. Abadi, R. L. Vieriu, S. Winkler, and N. Sebe, “ASCERTAIN: Emotion and personality recognition using commercial sensors,” *IEEE Trans. Affect. Comput.*, vol. 9, no. 2, pp. 147–160, Apr.–Jun. 2018.

[77] A. Ometov et al., “A survey on wearable technology: History, state-of-the-art and current challenges,” *Comput. Netw.*, vol. 193, 2021, Art. no. 108074.

[78] S. Lindner et al., “How do personality traits manifest in daily life of older adults?,” *Eur. J. Ageing*, vol. 19, no. 1, pp. 131–142, 2022.

[79] U. Ebner-Priemer, T. Kubiak, and K. Pawlik, “Ambulatory assessment,” *Eur. Psychol.*, vol. 14, no. 2, pp. 95–97, 2009.

[80] B. W. Nelson et al., “Guidelines for wrist-worn consumer wearable assessment of heart rate in biobehavioral research,” *NPJ Digit. Med.*, vol. 3, no. 1, 2020, Art. no. 90.

[81] R. J. Larsen and T. Ketelaar, “Personality and susceptibility to positive and negative emotional states,” *J. Pers. Social Psychol.,* vol. 61, no. 1, pp. 132–140, 1991.

[82] W. Han et al., “Mood states and everyday creativity: Employing an experience sampling method and a day reconstruction method,” *Front. Psychol.*, vol. 10, 2019, Art. no. 1698.

[83] R. E. Lucas et al., “A direct comparison of the day reconstruction method (DRM) and the experience sampling method (ESM),” *J. Pers. Social Psychol.*, vol. 120, no. 3, pp. 816–835, 2021.

[84] D. J. Miller, D. D. Vachon, and D. R. Lynam, “Neuroticism, negative affect, and negative affect instability: Establishing convergent and discriminant validity using ecological momentary assessment,” *Pers. Individual Differences*, vol. 47, no. 8, pp. 873–877, 2009.

[85] K. M. DeNeve and H. Cooper, “The happy personality: A meta-analysis of 137 personality traits and subjective well-being,” *Psychol. Bull.*, vol. 124, no. 2, pp. 197–229, 1998.

[86] P. Verduyn and K. Brans, “The relationship between extraversion, neuroticism and aspects of trait affect,” *Pers. Individual Differences*, vol. 52, no. 6, pp. 664–669, 2012.

[87] D. K. Darnell and P. A. Krieg, “Student engagement, assessed using heart rate, shows no reset following active learning sessions in lectures,” *PLoS One*, vol. 14, no. 12, 2019, Art. no. e0225709.

[88] L. Rösler et al., “Social anxiety is associated with heart rate but not gaze behavior in a real social interaction,” *J. Behav. Ther. Exp. Psychiatry*, vol. 70, 2021, Art. no. 101600.

[89] M. Ermes, J. PÄrkkÄ, J. MÄntyjÄrvi, and I. Korhonen, “Detection of daily activities and sports with wearable sensors in controlled and uncontrolled conditions,” *IEEE Trans. Inf. Technol. Biomed.*, vol. 12, no. 1, pp. 20–26, Jan. 2008.

[90] A. Darzi et al., “Classification of multiple psychological dimensions in computer game players using physiology, performance, and personality characteristics,” *Front. Neurosci.*, vol. 13, 2019, Art. no. 1278.

[91] Y. J. Weisberg, C. G. DeYoung, and J. B. Hirsh, “Gender differences in personality across the ten aspects of the big five,” *Front. Psychol.*, vol. 2, 2011, Art. no. 178.

[92] D. P. Schmitt et al., “Personality and gender differences in global perspective,” *Int. J. Psychol.*, vol. 52, pp. 45–56, 2017.

[93] S. Epstein, “The stability of behavior: I. On predicting most of the people much of the time,” *J. Pers. Social Psychol.*, vol. 37, no. 7, pp. 1097–1126, 1979.

[94] S. Epstein, “Aggregation and beyond: Some basic issues on the prediction of behavior,” *J. Pers.*, vol. 51, no. 3, pp. 360–392, 1983.

[95] A. Basu, A. Dasgupta, A. Thyagharajan, A. Routray, R. Guha, and P. Mitra, “A portable personality recognizer based on affective state classification using spectral fusion of features,” *IEEE Trans. Affect. Comput.*, vol. 9, no. 3, pp. 330–342, Jul.–Sep. 2018.

[96] Y. Qiao et al., “Multifunctional and high-performance electronic skin based on silver nanowires bridging graphene,” *Carbon*, vol. 156, pp. 253–260, 2020.

[97] S. Halem et al., “Moments that matter? On the complexity of using triggers based on skin conductance to sample arousing events within an experience sampling framework,” *Eur. J. Pers.*, vol. 34, pp. 794–807, 2020.

[98] A. J. Perez and S. Zeadally, “Recent advances in wearable sensing technologies,” *Sensors*, vol. 21, no. 20, 2021, Art. no. 6828.

[99] Y. Zhang et al., “Wearable neurophysiological recordings in middle-school classroom correlate with students’ academic performance,” *Front. Hum. Neurosci.*, vol. 12, 2018, Art. no. 457.

[100] Z. Qu, J. Chen, B. Li, J. Tan, D. Zhang, and Y. Zhang, “Measurement of high-school students’ trait math anxiety using neurophysiological recordings during math exam,” *IEEE Access*, vol. 8, pp. 57460–57471, 2020.

[101] K. de Barbaro, “Automated sensing of daily activity: A new lens into development,” *Devlop. Psychobiol.*, vol. 61, no. 3, pp. 444–464, 2019.

[102] Q. Ni, A. B. Garcia Hernando, and L. C. Ivan Pau de, “A context-aware system infrastructure for monitoring activities of daily living in smart home,” *J. Sensors*, vol. 2016, 2016, Art. no. 9493047.

[103] N. P. Rocha et al., “Systematic literature review of context-awareness applications supported by smart cities’ infrastructures,” *SN Appl. Sci.*, vol. 4, no. 4, 2022, Art. no. 90.

[104] N. Hegde, M. Bries, T. Swibas, E. Melanson, and E. Sazonov, “Automatic recognition of activities of daily living utilizing insole-based and wrist-worn wearable sensors,” *IEEE J. Biomed. Health Inform.*, vol. 22, no. 4, pp. 979–988, Jul. 2018.