ASCERTAIN: Emotion and Personality Recognition using Commercial Sensors

Ramanathan Subramanian, *Senior Member, IEEE,* Julia Wache *Student Member, IEEE,*
Mojtaba Khomami Abadi, *Student Member, IEEE,* Radu L. Vieriu, *Member, IEEE,* Stefan Winkler, *Senior Member, IEEE,* Nicu Sebe, *Senior Member, IEEE*

Abstract—We present ASCERTAIN– a multimodal databa**AS**e for impli**C**it p**ER**sonali**T**y and **A**ffect recognit**IO**N using commercial physiological sensors. To our knowledge, ASCERTAIN is the first database to connect *personality traits* and *emotional states* via *physiological responses*. ASCERTAIN contains big-five personality scales and emotional self-ratings of 58 users along with their Electroencephalogram (EEG), Electrocardiogram (ECG), Galvanic Skin Response (GSR) and facial activity data, recorded using off-the-shelf sensors while viewing affective movie clips. We first examine relationships between users' affective ratings and personality scales in the context of prior observations, and then study linear and non-linear physiological correlates of emotion and personality. Our analysis suggests that the emotion–personality relationship is better captured by non-linear rather than linear statistics. We finally attempt binary emotion and personality trait recognition using physiological features. Experimental results cumulatively confirm that personality differences are better revealed while comparing user responses to emotionally homogeneous videos, and above-chance recognition is achieved for both affective and personality dimensions.

Index Terms—Emotion and Personality recognition, Physiological signals, Multimodal analysis, Commercial sensors

✦

1 INTRODUCTION

DEspite rapid advances in Human-computer Interaction (HCI) and relentless endeavors to improve user experience with computer systems, the need for agents to *recognize* and *adapt* to the **affective state** of users has been widely acknowledged. While being a critical component of human behavior, affect is nevertheless a highly subjective phenomenon influenced by a number of contextual and psychological factors including **personality**.

The personality–affect relationship has been actively studied ever since a correlation between the two was proposed in Eysenck's personality model [1]. Eysenck posited that Extraversion, the personality dimension that describes a person as either talkative or reserved, is accompanied by low cortical arousal– *i.e.,* extraverts require more external stimulations than introverts. His model also proposed that neurotics, characterized by negative feelings such as depression and anxiety, are more sensitive to external stimulation and become easily upset or nervous due to minor stressors.

Many affective studies have attempted to validate and extend Eyesenk's findings. Some have employed explicit user feedback in the form of affective self-ratings [2], [3], while others have measured implicit user responses such as

Electroencephalogram (EEG) activity [4] and heart rate [5] for their analyses. However, few works have investigated affective correlates of traits other than Extraversion and Neuroticism. Conversely, social psychology studies have examined personality mainly via non-verbal social behavioral cues (see [6] for a review), but few works have modeled personality traits based on emotional behavior. Conducting studies to examine the personality–affect relationship is precluded by problems such as subject preparation time, invasiveness of sensing equipment and the paucity of reliable annotators for annotating emotional attributes.

This work builds on [7] and examines the influence of personality differences on users' affective behavior via the ASCERTAIN database[1]. We utilize ASCERTAIN to (i) understand the relation between *emotional attributes* and *personality traits*, and (ii) characterize both via users' physiological responses. ASCERTAIN contains personality scores and emotional self-ratings of 58 users in addition to their affective physiological responses. More specifically, ASCERTAIN is used to model users' emotional states and *big-five* personality traits via heart rate (Electrocardiogram or ECG), galvanic skin response (GSR), EEG and facial activity patterns observed while viewing 36 affective movie clips.

We specifically designed a study with movie scenes as they effectively evoke emotions [8], [9], as typified by genres such as *thriller*, *comedy* or *horror*. Also, different from existing affective databases such as DEAP [10], MAHNOB [11] and DECAF [9], ASCERTAIN comprises data recorded exclusively using commercial sensors to ensure ecological validity and scalability of the employed framework for large-scale profiling applications.

Using the ASCERTAIN data, we first examine correla-

- *Ramanathan Subramanian is with the International Institute of Information Technology, Hyderabad, India. (Email: s.ramanathan@iiit.ac.in)*
- *Julia Wache, Mojtaba Khomami Abadi, Radu L.Vieriu and Nicu Sebe are with the Dept. of Information Engineering and Computer Science, University of Trento, 38123 Trento, Italy. (Email: julia.wache@unitn.it, khomamiabadi@disi.unitn.it, radulaurentiu.vieriu@unitn.it, sebe@disi.unitn.it)*
- *Stefan Winkler is with the Advanced Digital Sciences Center, University of Illinois at Urbana-Champaign, Singapore. (Email: Stefan.Winkler@adsc.com.sg)*

1. http://mhug.disi.unitn.it/index.php/datasets/ascertain/

tions among users' valence (V) and arousal (A) self-ratings and their personality dimensions. We then attempt to isolate physiological correlates of emotion and personality. Our analyses suggest that the relationships among emotional attributes and personality traits are better captured by non-linear rather than linear statistics. Finally, we present single-trial (binary) recognition of A,V and the big-five traits considering physiological responses observed over (a) *all*, and (b) *emotionally homogeneous* (*e.g.,* high A, high V) clips. Superior personality recognition is achieved for (b), implying that personality differences are better revealed by comparing responses to emotionally similar stimuli. The salient aspects of ASCERTAIN are:

1. To our knowledge, ASCERTAIN is the first physiological database that facilitates both emotion and personality recognition. In social psychology, personality traits are routinely modeled via questionnaires or social behavioral cues. Instead, this is one of the first works to assess personality traits via affective physiological responses (the only other work to this end is [12]).

2. Different from the DEAP [10], MAHNOB [11] and DECAF [9] databases, we use *wearable, off-the-shelf* sensors for physiological recordings. This enhances the ecological validity of the ASCERTAIN framework, and above-chance recognition of emotion and personality affirms its utility and promise for commercial applications.

3. We present interesting insights concerning correlations among affective and personality attributes. Our analyses suggest that the emotion–personality relationship is better captured via non-linear statistics. Also, personality differences are better revealed by comparing user responses to emotionally similar videos (or more generally, under similar affect inducement).

From here on, Section 2 reviews related literature to motivate the need for ASCERTAIN, while Section 3 details the materials and methods employed for data compilation. Section 4 presents descriptive statistics, while correlations among users' affective ratings and personality dimensions are analyzed in Section 5. Section 6 details physiological correlates of emotion and personality, while Section 7 presents recognition experiments. Section 8 discusses the correlation and recognition results, and Section 9 concludes the paper.

2 RELATED WORK

This section reviews related work focusing on (a) multimodal affect recognition, (b) personality assessment and (c) the personality–affect relationship.

2.1 Multimodal affect recognition

As emotions are conveyed by content creators using multiple means (audio, video), and expressed by humans in a number of ways (facial expressions, speech and physiological responses), many affect recognition (AR) methods employ a multimodal framework. Common content-based modalities employed for AR include audio [14]–[16], visual [17]–[19] and audio-visual [20]–[22]. Recent AR methodologies have focused on monitoring user behavior via the use of physiological sensors (see [23] for a review). Emotions induced by music clips are recognized via heart rate, muscle movements, skin conductivity and respiration changes

in [24]. Lisetti *et al.* [25] use GSR, heart rate and temperature signals to recognize emotional states. As part of the HUMAINE project [13], three naturalistic and six induced affective databases containing multimodal data (including physiological signals) are compiled from 8–125 participants. Tavakoli *et al.* [26] examine the utility of various eye fixation and saccade-based features for valence recognition, while Subramanian *et al.* [27] correlate user responses with eye movement patterns to discuss the impact of emotions on visual attention and memory.

Koelstra *et al.* [10] analyze blood volume pressure, respiration rate, skin temperature and Electrooculogram (EOG) patterns for recognizing emotional states induced by 40 music videos. MAHNOB-HCI [11] is a multimodal database containing synchronized face video, speech, eye-gaze and physiological recordings from 27 users. Abadi *et al.* [9] study Magnetoencephalogram (MEG), Electromyogram (EMG), EOG and ECG responses from users for music and movie clips, and conclude that better emotion elicitation and AR are achieved with movie clips.

2.2 Personality recognition

The *big-five* or five-factor model [28] describes human personality in terms of five dimensions– Extraversion (*sociable* vs *reserved*), Neuroticism or the degree of emotional stability (*nervous* vs *confident*), Agreeableness (*compassionate* vs *dispassionate*), Conscientiousness (*dutiful* vs *easy-going*) and Openness (*curious/creative* vs *cautious/conservative*).

A comprehensive survey of personality computing approaches is presented in [6]. The traditional means to model personality traits are questionnaires or self-reports. Argamon *et al.* [29] use lexical cues from informal texts for recognizing Extraversion (*Ex*) and Neuroticism (*Neu*). Olguin *et al.* [30] and Alameda-Pineda *et al.* [31] show that non-verbal behavioral measures acquired using a sociometric badge such as the amount of speech and physical activity, number of face-to-face interactions and physical proximity to other objects is highly correlated with personality. Much work has since employed non-verbal behavioral cues in social settings for personality recognition including [32], where *Ex* is recognized using speech and social attention cues in round-table meetings, while [33], [34] predict *Ex* and *Neu* from proxemic and attention cues in party settings.

Among works that have attempted recognition of all five personality factors, Mairesse *et al.* [35] use acoustic and lexical features, while Staiano *et al.* [36] analyze structural features of individuals' social networks. Srivastava *et al.* [37] automatically complete personality questionnaires for 50 movie characters utilizing lexical, audio and visual behavioral cues. Brouwer *et al.* [38] estimate personality traits via physiological measures, which are revealed sub-consciously and more genuinely (less prone to manipulation) than questionnaire answers. In a gaming-based study, they observe a negative correlation between (i) heart rate and *Ex*, and (ii) skin-conductance and *Neu*.

2.3 Personality-Affect relationship

The relationship between personality and affect has been extensively examined in social psychology [39], but not in a computational setting. Eysenck's seminal personality

TABLE 1
Comparison of user-centered affective databases. 'var' denotes variable.

Name	No. subjects	No. stimuli	Recorded signals	Annotations		Comments
				Affect	Personality	
HUMAINE [13]	var	var	audio, visual, physiological	yes	no	includes 6 sub-collections (some non-public)
DEAP [10]	32	40	physiological	yes	no	focus on music videos
DECAF [9]	30	76	face, physiological	yes	no	compares music and movie clips
MAHNOB-HCI [11]	27	20	face, audio, eye gaze, physiological	yes	no	includes video and image stimuli
ASCERTAIN	**58**	**36**	face, physiological	yes	yes	connects emotion and personality

theory [1] posits that extraverts require more external stimulation than introverts, and that neurotics are aroused more easily. Many studies have since studied the personality–affect relationship by examining explicit or implicit user responses. Personality effects on brain activation related to valence (V) and arousal (A) is investigated in [3], which concludes that *Neu* correlates negatively with positive V, and positively with A. In an EEG-based study [4], a negative correlation is observed between *Ex* and A, while a positive correlation is noted between *Neu* and A especially for negative valence stimuli.

The impact of personality traits on affective user ratings is studied using path analysis in [40]. Feedback scores from 133 students are analyzed in [2] to conclude that neurotics experience positive emotions similar to emotionally stable counterparts in pleasant situations, even though they may experience negative emotions more strongly. Event-related potentials and heart rate changes are studied in [5] to confirm a positive correlation between *Neu* and A for negative stimuli, while a signal-detection task is used in [41] to suggest that extraverts are generally less aroused than introverts. Brumbaugh *et al.* [42] examine correlations among the big-five traits, and find *Ex* and *Neu* to be associated with increased A while viewing negative videos. Abadi *et al.* [12] attempt recognition of the big-five traits from affective physiological responses, and our work is most similar to theirs in this respect. Nevertheless, we consider more users and a larger stimulus set in this work (58 users and 36 clips vs 36 users and 16 clips in [12]), and show superior personality trait recognition on comparing physiological responses to emotionally homogeneous clips.

2.4 Spotting the research gap

Examination of related literature reveals that AR methodologies are increasingly becoming *user-centric* instead of *content-centric*, suggesting that emotions better manifest via human behavioral cues rather than multimedia content-based (typically audio, visual and speech-based) cues. Nevertheless, the influence of psychological factors such as personality on emotional behavior has hardly been examined, in spite of prior work suggesting that personality affects one's a) feelings [39], [43], b) emotional perception [3], [4] and c) multimedia preferences [44], [45].

Motivated by the above findings and the lack of publicly available data sets positioned at the intersection of personality and affect, we introduce ASCERTAIN, a multimodal corpus containing physiological recordings of users viewing emotional videos. ASCERTAIN allows for inferring both personality traits and emotional states from physiological signals. We record GSR, EEG, ECG signals using wearable sensors, and facial landmark trajectories (EMO) using a web-camera. In the light of recent technological

TABLE 2
Summary of the ASCERTAIN database.

Number of Participants	58
Number of Videos	36
Video Length	51–128 seconds ($\mu \pm \sigma$ = 80 $\pm$ 20)
Self-reported ratings	Arousal, Valence, Engagement Liking, Familiarity
Personality Scales	Extraversion, Agreeableness Conscientiousness, Neuroticism, Openness
Physiological signals	ECG, GSR, Frontal EEG, Facial features

developments, these signals can be acquired and analyzed instantaneously. Also, Wang and Ji [23] advocate the need for less-intrusive sensors to elicit natural emotional behavior from users. Use of wearable sensors is critical to ensure the ecological validity, repeatability and scalability of affective computing studies, which are typically conducted in controlled lab conditions and with small user groups.

Table 1 presents an overview of publicly available user-centric AR datasets. Apart from being one of the largest datasets in terms of the number of participants and stimuli examined for analysis, ASCERTAIN is also the first database to facilitate study of the personality–affect relationship.

3 ASCERTAIN OVERVIEW

Fig. 1 presents an overview of the ASCERTAIN framework and a summary of the compiled data is provided in Table 2. To study the personality–affect relationship, we recorded users' physiological responses as they viewed the affective movie clips used in [9]. Additionally, their explicit feedback, in the form of *arousal*, *valence*, *liking*, *engagement* and *familiarity* ratings, were obtained on viewing each clip. Finally, personality measures for the big-five dimensions were also compiled using a big-five marker scale (BFMS) questionnaire [46]. We now describe (1) the procedure adopted to compile users' emotional ratings, personality measures and physiological responses, and (2) the physiological features extracted to measure users' emotional responses.

3.1 Materials and Methods

Subjects: 58 university students (21 female, mean age = 30) participated in the study. All subjects were fluent in English and were habitual Hollywood movie watchers.

Materials: One PC with two monitors was used for the experiment. One monitor was used for video clip presentation at 1024×768 pixel resolution with 60 Hz screen refresh rate, and was placed roughly one meter before the user. The other monitor allowed the experimenter to verify the recorded sensor data. Following informed consent, physiological sensors were positioned on the user's body as shown in Fig. 2(a). The GSR sensor was tied
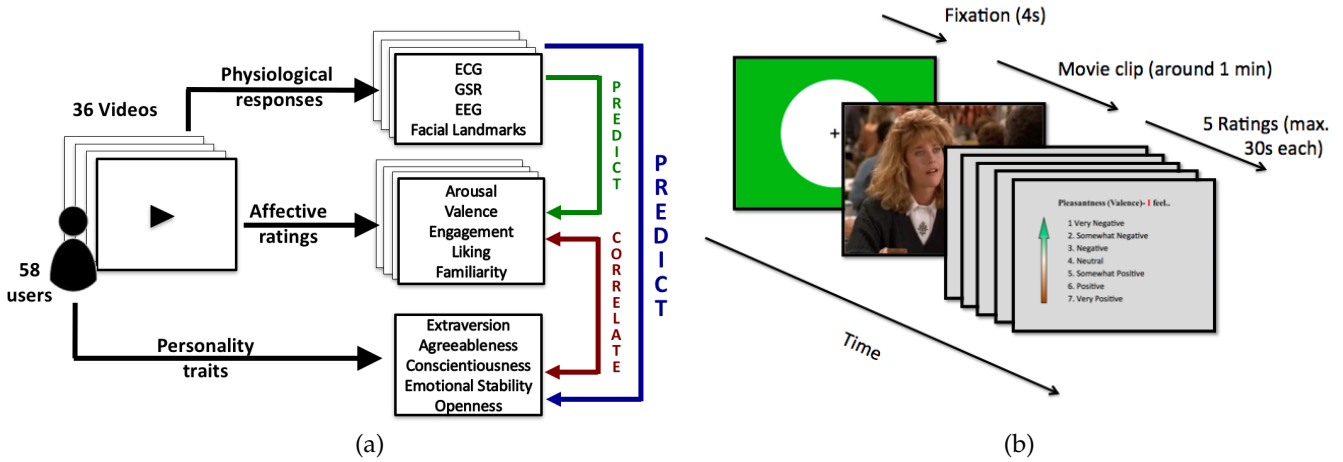
Fig. 1. (a) ASCERTAIN study overview. (b) Timeline for each trial.

to the left wrist, and two electrodes were fixed to the index and middle finger phalanges. Two measuring electrodes for ECG were placed at each arm crook, with the reference electrode placed at the left foot. A single dry-electrode EEG device was placed on the head like a normal headset, with the EEG sensor touching the forehead and the reference electrode clipped to the left ear. EEG data samples were logged using the `Lucid Scribe` software, and all sensor data were recorded via bluetooth. A webcam was used to record facial activity. Synchronized data recording and pre-processing were performed using MATLAB Psychtoolbox (http://psychtoolbox.org/).

Protocol: Each user performed the experiment in a session lasting about 90 minutes. Viewing of each movie clip is denoted as a *trial*. After two practice trials involving clips that were not part of the actual study, users watched movie clips randomly shown in two blocks of 18 trials, with a short break in-between to avoid fatigue. In each trial (Fig. 1(b)), a fixation cross was displayed for four seconds followed by clip presentation. After viewing each clip, users self-reported their emotional state in the form of affective ratings within a time limit of 30 seconds. They also completed a personality questionnaire after the experiment.

Stimuli: We adopted the 36 movie clips used in [9] for our study. These clips are between 51–127 s long (μ =80,

σ =20), and are shown to be uniformly distributed (9 clips per quadrant) over the arousal-valence (AV) plane.

Affective ratings: For each movie clip, we compiled valence (V) and arousal (A) ratings reflecting the user's affective impression. A 7-point scale was used with a -3 (*very negative*) to 3 (*very positive*) scale for V, and a 0 (*very boring*) to 6 (*very exciting*) scale for A. Likewise, ratings concerning engagement (*Did not pay attention – Totally attentive*), liking (*I hated it – I loved it*) and familiarity (*Never seen it before – Remember it very well*) were also acquired. Mean user V,A ratings for the 36 clips are plotted in Fig. 2(b), and are color-coded based on the ground-truth ratings from [9]. Ratings form a 'C'-shape in the AV plane, consistent with prior affective studies [9], [10].

Personality scores: Participants also completed the big-five marker scale (BFMS) questionnaire [46] which has been used in many personality recognition works [32]–[34]. Scale distributions for the big-five traits are shown in Fig. 2(c). The most and least variance in personality scores are noted for the Extraversion and Openness traits respectively.

3.2 Physiological feature extraction

We extracted physiological features corresponding to each trial over the final 50 seconds of stimulus presentation, owing to two reasons: (1) The clips used in [9] are not
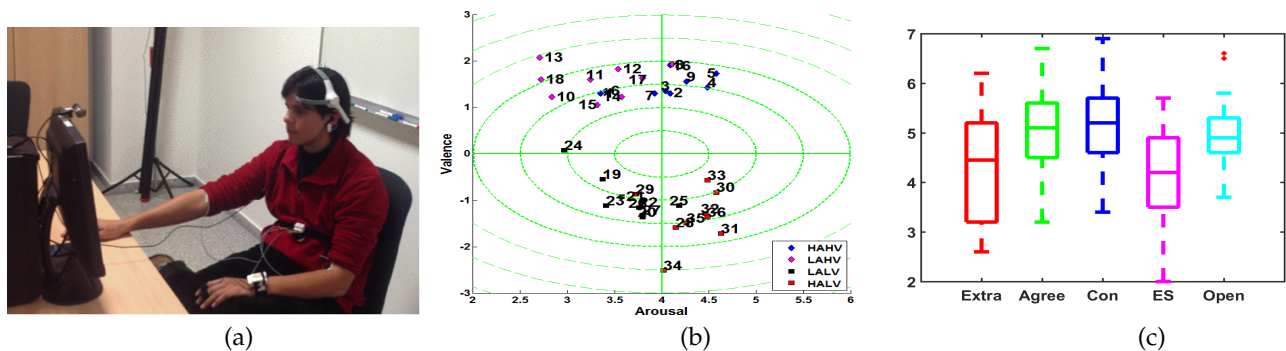

Fig. 2. (a) Participant with sensors (EEG, ECG and GSR visible) during the experiment, (b) Mean Arousal-Valence (AV) ratings for the 36 movie clips used in our experiment and (c) Box-plots showing distribution of the big-five personality trait scores for 58 users.

TABLE 3
Extracted features for each modality (feature dimension stated in parenthesis). *Statistics* denote mean, standard deviation (std), skewness, kurtosis of the raw feature over time, and % of times the feature value is above/below mean±std.

Modality	Extracted features
ECG (32)	Ten low frequency ([0-2.4] Hz) power spectral densities (PSDs), four very slow response ([0-0.04] Hz) PSDs, IBI, HR and HRV statistics.
GSR (31)	Mean skin resistance and mean of derivative, mean differential for negative values only (mean decrease rate during decay time), proportion of negative derivative samples, number of local minima in the GSR signal, average rising time of the GSR signal, spectral power in the [0-2.4] Hz band, zero crossing rate of skin conductance slow response ([0-0.2] Hz), zero crossing rate of skin conductance very slow response ([0-0.08] Hz), mean SCSR and SCVSR peak magnitude.
Frontal EEG (88)	Average of first derivative, proportion of negative differential samples, mean number of peaks, mean derivative of the inverse channel signal, average number of peaks in the inverse signal, statistics over each of the 8 signal channels provided by the Neurosky software.
EMO (72)	Statistics concerning horizontal and vertical movement of 12 motion units (MUs) specified in [47].

emotionally homogeneous, but are more emotional towards the end. (2) Some employed features (see Table 3) are nonlinear functions of the input signal length, and fixed time-intervals needed to be considered as the movie clips were of varying lengths. Descriptions of the physiological signals examined in this work are as follows.

Galvanic Skin Response (GSR): GSR measures transpiration rate of the skin. When two electrodes are positioned on the middle and index finger phalanges and a small current is sent through the body, resistance to current flow changes with the skin transpiration rate. Most of the GSR information is contained in low-frequency components, and the signal is recorded at 100 Hz sampling frequency with a commercial bluetooth sensor. Following [10], [11], [24], we extracted 31 GSR features listed in Table 3.

Electroencephalography (EEG): EEG measures small changes in the skull's electrical field produced by neural activity, and information is encoded in the EEG signal amplitude as well as in certain frequency components. We used a commercial, single dry-electrode EEG sensor[2], which records eight information channels sampled at 32 Hz. The recorded information includes frontal lobe activity, level of facial activation, eye-blink rate and strength, which are relevant emotional responses.

Electrocardiogram (ECG): Heart rate characteristics have been routinely used for user-centered emotion recognition. We performed R-peak detection on the ECG signal to compute users' inter-beat intervals (IBI), heart rate (HR), and the heart rate variability (HRV). We also extracted power spectral density (PSD) in low frequency bands as in [11], [24].

Facial landmark trajectories (EMO): A facial feature tracker [47] was used to compute displacements of 12 interest points or motion units (MU) in each video frame. We calculated 6 statistical measures for each landmark to obtain a total of 72 features (Table 3).

2. www.neurosky.com

3.3 Data Quality

A unique aspect of ASCERTAIN with respect to prior affective databases is that physiological signals are recorded using commercial and minimally invasive sensors that allow body movement of participants. However, it is well known that body movements can degrade quality of the recorded data, and such degradation may be difficult to detect using automated methods. Therefore, we plotted the recorded data for each modality and trial, and rated the data quality manually on a scale of 1 (*good data*)–5 (*missing data*). For ECG, we evaluated the raw signal from each arm as well as the R-peak amplitude. The presence/absence of facial tracks and correctness of the tracked facial locations were noted for EMO. For GSR, we examined the extent of data noise, and rated EEG (i) on the raw signal, (ii) by summarizing the quality of δ (< 4 Hz), θ (4–7 Hz), α (8–15 Hz), β (16–31 Hz) and γ (> 31 Hz) frequency bands, and (iii) on the pre-calculated *attention* and *meditation* channels available as part of the EEG data. Plots and tables with explanations on data quality are available with the dataset. Fig. 3 presents an overview of the data quality for the four considered modalities, with the proportion of trials for which the quality varies from 1–5 highlighted. About 70% of the recorded data is good (corresponding to levels 1-3) for all modalities except ECG, with GSR data being the cleanest. Maximum missing data is noted for EEG, reflecting the sensitivity of the EEG device to head movements.
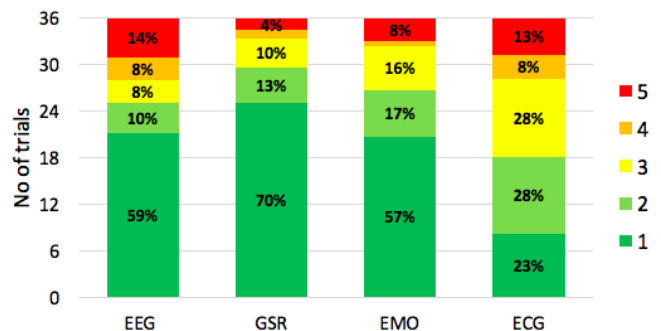

Fig. 3. Bar plot showing proportion of trials for which data quality ranges from best (1) to worst (5).

4 DESCRIPTIVE STATISTICS

In this section, we present statistics relating to user self-reports and personality scores.

4.1 Analysis of Self-ratings

As mentioned previously, we selected 36 movie clips such that their emotional ratings were distributed uniformly over the AV plane as per ground-truth ratings in [9], with 9 clips each corresponding to the HAHV (high arousal-high valence), LAHV (low arousal-high valence), LALV (low arousal-low valence) and HALV (high arousal-low valence) quadrants[3]. The targeted affective state was mostly reached during the ASCERTAIN study as shown in Fig. 2(b). A two-sample t-test revealed significantly higher mean A ratings

3. For consistency's sake, quadrant-wise video labels derived based on ratings from [9] are used in this work.

for HA clips as compared to LA clips ($t(34) = 5.1253, p < 0.0001$). Similarly, mean V ratings for HV and LV clips were significantly different ($t(34) = 17.6613, p < 0.00005$). Overall, emotion elicitation was more consistent for valence as in prior works [9], [10].

We computed agreement among participants' A,V ratings using the Krippendorff's alpha metric– agreement for A and V were respectively found to be 0.12 and 0.58, implying more consensus for clip valence as above. We then computed the agreement between the ASCERTAIN and DECAF [9] populations using the Cohen's Kappa (κ) measure. To this end, we computed κ between ground-truth (GT) labels from [9] and each user's A,V labels assigned as *high/low* based on the mean rating– the mean agreement over all users for A and V was found to be 0.24 and 0.73 respectively. We also computed the κ measure between GT and the ASCERTAIN population based on the mean A,V rating of all users– here, an agreement of 0.39 was observed for A and 0.61 for V. Overall, these measures suggest that while individual-level differences exist in affective perception of the movie clips, there is moderate to substantial agreement between assessments of the ASCERTAIN and DECAF populations implying that the considered movie clips are effective for emotion elicitation.

Fig. 4 presents box-plots describing the distribution of the arousal (A), valence (V), engagement (E), liking (L) and familiarity (F) user ratings for (i) all, and (ii) quadrant-based videos. Clearly, low-arousal videos are perceived as more 'neutral' in terms of A and V, which leads to the 'C' shape in Fig. 2(b). All videos are perceived as sufficiently engaging, while HV clips are evidently more liked than LV clips. Also, the presented movie clips were not very conversant to participants, suggesting that the ASCERTAIN findings are overall unlikely to be influenced by familiarity biases.

4.2 Affective Ratings vs Personality Scales

To examine relationships between the different user ratings, we computed Pearson correlations among self-reported attributes as shown in Table 4. Since the analysis involves attribute ratings provided by 58 users for each of the 36 clips, we accounted for multiple comparisons by limiting the false discovery rate (FDR) to within 5% using the procedure outlined in [48]. Highlighted numbers denote correlations found to be significant over at least 15 users (25% of the population) adopting the above methodology.

Focusing on significant correlations, A is moderately correlated with E, while V is found to correlate strongly with L mirroring the observations of Koelstra *et al.* [10]. A moderate and significant correlation is noted between E and L implying that engaging videos are likely to appeal to viewers' senses, and similarly, between F and L confirming the mere exposure effect observed in [49] attributing liking to familiarity. Nevertheless, different from [10] with music videos where a moderate correlation is noted between A and V ratings, we notice that the A and V dimensions are uncorrelated for the ASCERTAIN study, which again reinforces the utility of movie clips as good control stimuli. To validate our experimental design, we tested for effects of video length on A,V ratings but did not find any.

Table 5 presents Pearson correlations between personality dimensions. Again focusing on significant correla-

TABLE 4
Mean Pearson correlations between self-ratings across users. *s denote significant correlations ($p < 0.05$) upon limiting FDR to 5%.

	A	V	E	L	F
Arousal	1	0.02	0.42*	0.19	0.15
Valence		1	0.21	0.68*	0.17
Engagement			1	0.42*	0.24
Liking				1	0.34*
Familiarity					1

TABLE 5
Pearson correlations between personality dimensions ($* \Rightarrow p < 0.05$)

	E	A	Co	ES	O
Extraversion	1	0.36*	0.19	-0.12	0.30*
Agreeableness		1	0.21	0.34*	0.30*
Conscientiousness			1	0.26	0.04
Emotional Stability				1	-0.10
Openness					1

TABLE 6
Partial correlations between personality scales and self-ratings ($* \Rightarrow p < 0.05$).

		Ex	Ag	Co	ES	O
All	Arousal	0.03	-0.10	0.05	0.07	0.06
	Valence	0.19	-0.02	0.02	-0.18	0.07
	Engage	**-0.30***	0.01	0.09	0.00	-0.10
	Liking	-0.13	0.07	-0.22	0.21	-0.02
HAHV	Arousal	-0.01	0.02	-0.01	0.22	-0.11
	Valence	-0.12	**-0.38***	-0.11	-0.12	-0.16
	Engage	**-0.30***	0.16	0.17	0.10	-0.09
	Liking	0.20	0.22	-0.00	0.10	0.22
LAHV	Arousal	-0.06	0.07	0.06	0.11	0.04
	Valence	-0.03	0.05	-0.08	-0.10	0.23
	Engage	-0.22	0.03	-0.06	-0.10	-0.24
	Liking	0.20	-0.01	0.13	0.17	0.12
LALV	Arousal	0.03	-0.09	0.02	-0.07	0.09
	Valence	0.20	0.06	0.01	-0.22	0.15
	Engage	-0.22	-0.01	0.02	-0.05	-0.04
	Liking	-0.14	0.03	-0.19	0.12	-0.10
HALV	Arousal	0.20	-0.25	-0.00	-0.16	0.09
	Valence	0.22	0.01	0.09	-0.06	-0.05
	Engage	**-0.30***	0.01	0.10	0.12	-0.10
	Liking	**-0.26***	0.03	**-0.35***	0.12	-0.07

TABLE 7
R^2 and best three predictors for the five personality dimensions. Full model coefficients are shown in parentheses. $* \Rightarrow p < 0.05$.

	Ex	Ag	Co	ES	O
All	**0.14*** (0.14)	0.02 (0.02)	0.07 (0.07)	0.06 (0.06)	0.02 (0.02)
	V,E,L	A,V,L	V,E,L	A,V,L	A,V,E
HAHV	**0.16*** (0.16)	**0.17*** (0.17)	0.05 (0.05)	**0.12*** (0.13)	0.05 (0.06)
	V,E,L	V,E,L	A,V,E	A,V,L	A,V,L
LAHV	0.07 (0.08)	0.02 (0.02)	0.02 (0.02)	0.04 (0.05)	**0.13*** (0.13)
	A,E,L	A,V,E	A,E,L	A,E,L	V,E,L
LALV	0.12 (0.12)	0.02 (0.02)	0.05 (0.05)	0.05 (0.05)	0.03 (0.03)
	V,E,L	A,V,L	A,E,L	A,V,L	A,V,L
HALV	**0.16*** (0.20)	0.09 (0.09)	0.15 (0.16)	0.04 (0.05)	0.03 (0.04)
	V,E,L	A,V,L	V,E,L	A,E,L	A,E,L

tions, moderate and positive correlations are noted between Extraversion (*Ex*) and Agreeableness (*Ag*), as well as between *Ex* and Openness (*O*)– prior studies have noted that *Ex* and *O* are correlated via the sensation seeking construct [50]. *Ag* is also found to moderately and positively correlate with Emotional Stability (*ES*) and *O*. Conversely, weakly negative-but-insignificant correlations are observed between (i) *Ex* and *ES*, and (ii) *ES* and *O*.

Partial correlations between emotional and personality attributes are tabulated in Table 6. Considering all movie clips, a significant and moderately negative correlation is noted between *Ex* and E, implying that introverts were more immersed with emotional clips during the movie-watching
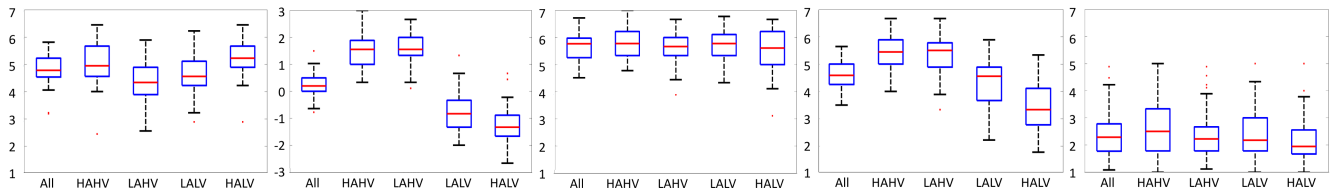
Fig. 4. Boxplots of the mean Arousal, Valence, Engagement, Liking and Familiarity ratings for the different video sets.

task. A few more significant correlates are observed when mean ratings for quadrant-wise (or emotionally similar) videos are considered. Delineating, *Ag* is negatively correlated with V for HAHV videos, while the negative correlation between *Ex* and E manifests for high-arousal (HAHV and HALV) stimuli. Also notable is the moderately negative correlation between *Ex* and L, and also between Conscientiousness and L for HALV movie clips.

We also performed linear regression analyses with user self ratings as predictors and personality attributes as the target variables for the different video sets, and the coefficients of determination/squared correlations (R^2) for the different video sets are presented in Table 7. R^2 values with the three best predictors along with the predictor names are listed outside parentheses, while squared correlations with the full model are listed within braces. Considering all movie clips, the best linear model is obtained for *Ex* with V, E and L ratings as predictors. Among the four AV quadrants, significant squared correlations are observed for the *Ex* and *Ag* traits with V,E,L predictors, and for the *ES* trait with arousal, valence and liking ratings as predictors for HAHV clips. A significant model is also obtained for Openness with V,E,L predictors considering mildly positive HALV clips. Overall, it is easy to observe from the table that (i) there is little difference in the predictive power of the best-three-predictor and full models, and (ii) the linear models have rather limited predictive power, with the best model explaining only 17% of the personality scale variance. Cumulatively, Tables 6 and 7 cumulatively suggest that the relationship between emotional and personality variables is not well modeled using linear statistics, and it is perhaps worthwhile to explore the use of non-linear measures to this end. From here on, given the high degree of correlation between A and E and between the V and L, we will only focus on A and V dimensions in the rest of the paper.

4.3 Mutual Information Analysis

Mutual information (MI) is a popular metric to capture non-linear relationships between two random variables, and measures how much information is known about one variable given the other. Formally, the MI between two random vectors $X = \{x\}$ and $Y = \{y\}$ is defined as:
$MI(X,Y) = \sum_{x,y} P_{XY}(x,y) log \frac{P_{XY}(x,y)}{P_X(x).P_Y(y)}$ where $p_{XY}(x,y)$ is the joint probability distribution, while $P_X(x)$ and $P_Y(y)$ are the respective marginal probabilities. We attempted to describe the relationship between emotional ratings and personality scales via the normalized mutual information (NMI) index [51] defined as: $NMI(X,Y) = \frac{MI(X,Y)}{\sqrt{(H(X)H(Y))}}$, where $H(X)$ and $H(Y)$ denote entropies of X and Y.

NMI with personality scales for arousal and valence ratings are shown in Fig. 5. In contrast to linear correlations,

both A and V share a high degree of mutual information with all five personality traits. Considering all movie clips, emotional ratings share slightly higher MI with A than with V. Also, a strictly higher MI measure is noted when emotionally similar clips are considered instead of all clips. Among personality traits, *Ex* and Conscientiousness (*Con*) share the most MI with V,A attributes– in contrast, little correlation is observed between *Con* and A,V in Table 6). Conversely, lowest MI is noted for Openness (*O*). One notable difference exists between A and V though– higher MI with arousal is noted for high HV clips, while for all personality traits barring *Ag*, greater MI with valence is observed for LV clips than for HV clips.
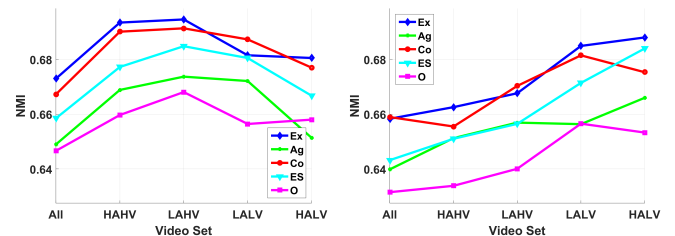

Fig. 5. NMI between big-five trait scales and A (left), V (right) ratings.

5 PERSONALITY MEASURES VS USER RATINGS

We now examine the relationship between user V,A ratings and personality scales in the context of hypotheses (H1–H3) put forth in the literature. To this end, we determined *high/low* trait groups (*e.g.*, emotional stable vs neurotic) for each personality dimension by dichotomizing personality measures based on the median score– this generated balanced *high* and *low* sets for the *Ex* and *ES* traits, and an unbalanced split for the remaining traits, with the most imbalance (33 vs 25) noted for *Ag*. We then proceeded to analyze the affective ratings for each group.
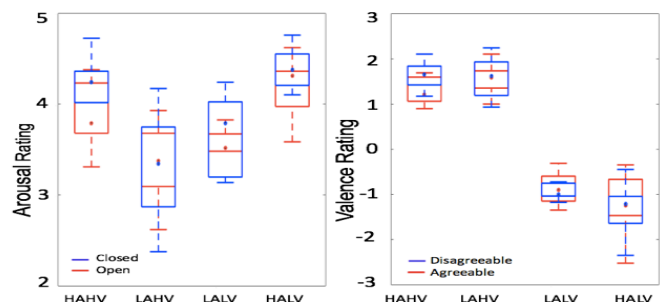

Fig. 6. Quadrant-wise comparisons of (left) A ratings by *open* and *closed* groups, and (right) V ratings by *agreeable* and *disagreeable* groups .

5.1 H1: Extraversion vs Arousal and Valence

The correlation between Extraversion and arousal has been investigated in many studies– EEG measurements [4], signal detection analysis [41] and fMRI [3] have shown lower arousal in extraverts as compared to introverts, consistent with Eyesenck's personality theory. Also, *Ex* has been found to correlate with positive valence in a number of works [52]. Analyses presented in Table 6 reveal little correlation between *Ex* and A for all video categories. While two-tailed *t*-tests confirmed that extraverts and introverts rated high A and low A videos differently ($p < 0.00001$ in both cases), no differences could be identified between their A ratings excepting that extraverts provided marginally lower ratings for HA clips ($t(56) = -1.4423, p = 0.0774$, left-tailed). Focusing on V ratings, positive correlation between *Ex* and V breaks down for HV clips in Table 6. Two-sample *t*-tests also failed to reveal any differences. Therefore, statistical analyses weakly support the negative correlation between *Ex* and A, but do not corroborate the positive correlation between *Ex* and V.

5.2 H2: Neuroticism vs Arousal and Valence

The relationship between *Neu* and A has also been extensively studied– a positive correlation between *Neu* and A is revealed through fMRI responses in [3], and EEG analysis [4] corroborates this observation for negative V stimuli. [2] further remarks that neurotics experience negative emotions stronger than emotionally stable persons. In contrast, differing observations have been made regarding the relationship between *Neu* and V. Negative correlation between *Neu* and positive V is noted in [3], whereas a positive relationship between the two for low A stimuli is observed in [40]. [2] remarks that the *Neu*-V relation is moderated by situation– while neurotics may feel less positive in unpleasant situations, they experience positive emotions as strongly as *ES* counterparts in pleasant conditions.

Negative correlation between Emotional Stability (*ES*) and A (or positive correlation between *Neu* and A) is noted only for HALV clips in Table 6. However, post-hoc *t*-tests failed to reveal differences between A ratings for the two categories. Also, Table 6 generally suggests a negative correlation between *ES* and V– *t*-test comparisons further revealed marginally lower V ratings provided by *ES* subjects for LALV clips ($t(16) = -1.3712, p = 0.0946$, left-tailed). Overall, our data does not support the positive relationship between *Neu* and A, and suggests a weakly positive correlation between *Neu* and V.

5.3 H3: Openness vs Valence and Arousal

Among the few works to study Openness, [40] notes a positive correlation between Openness (*O*) and V under low arousal conditions, which is attributed to the intelligence and sensitivity of creative individuals[4], enabling them to better appreciate subtly emotional stimuli. Table 6 echoes a positive (even if insignificant) correlation between *O* and V for LA clips but post-hoc *t*-tests to compare V ratings of *open* and *closed* groups failed to reveal any differences. However, we noted that *closed* individuals felt

4. Creativity strongly correlates with Openness [53].

somewhat more aroused by HA clips than *open* individuals ($t(56) = -1.5011, p = 0.0695$, left-tailed) as shown in Fig. 6(a). Fine-grained analysis via left-tailed *t*-tests to compare quadrant-wise ratings again revealed the slightly higher arousal experienced by *closed* subjects for HAHV clips ($t(16) = -1.3753, p = 0.0940$, left-tailed). In summary, our data weakly confirms a positive relationship between *O* and V as noted in [40], but suggests a negative correlation between *O* and A.

5.4 Agreeableness and Conscientiousness

Table 6 shows a negative but insignificant correlation between *Ag* and A for HALV videos. Comparison of A ratings by *agreeable* and *disagreeable* groups revealed marginally lower A for *agreeable* subjects for HA clips ($t(56) = -1.2964, p = 0.10$, left-tailed), and subsequent quadrant-wise comparisons attributed this finding to significantly lower A ratings provided by the *agreeable* group for strongly negative HALV clips ($t(16) = -2.6587, p < 0.01$, left-tailed). This trend could possibly be attributed to the association of *disagreeable* persons with negative feelings such as deceit and suspicion. Table 6 also shows a negative correlation between *Ag* and V for highly positive HAHV clips. *T*-test comparisons again revealed that *agreeable* subjects provided somewhat lower V ratings for HV clips ($t(56) = -1.4285, p = 0.0793$, left-tailed), and this was particularly true of HAHV clips for which significantly lower ratings were provided by the *agreeable* group ($t(16) = -2.0878, p < 0.05$, left-tailed). Conscientiousness scale differences did not influence VA ratings in any way.

6 Physiological correlates of emotion and personality

Linear and non-linear analyses presented in the previous sections suggest that correlations between emotional and personality attributes are better revealed while examining user responses to emotionally similar clips. If explicit ratings provided by users are a conscious reflection of their emotional perception, then the analyses employing physiological signals should also reveal similar patterns. We attempt to identify linear and non-linear physiological correlates of emotion and personality considering responses to *all* and *quadrant-specific* clips in this section.

6.1 Linear correlates of Emotion and Personality

We attempted to discover physiological correlates of emotional and the big-five personality attributes via partial Pearson correlations. Given the large number of extracted physiological features (Table 3) as compared to the population size for this study, we first performed a principal component analysis (PCA) on each feature modality to avoid overfitting, and retained those components that explained 99% of the variance. This gave us 8–9 predictors for each of the considered modalities. Table 8 presents correlations between these principal components, users' affective ratings and personality scales (R° denotes number of significant correlates). For affective dimensions, we determined significant correlates considering mean user V,A ratings provided for the 36 clips. We also trained regression models with

TABLE 8
Physiological correlates of emotion and personality attributes. R° denotes the number of significant feature correlates, while R^2 is the coefficient of determination for the regression model with the significant correlates as predictors. Bold values denote linear regression models with a significant R^2 statistic.

Video Set	Feature	Arousal R^o	R^2	Valence R^o	R^2	Extra. R^o	R^2	Agreeable R^o	R^2	Conscient R^o	R^2	Em. Stab. R^o	R^2	Open R^o	R^2
All	ECG	1	0.25			1	0.25			1	0.32	2	0.30	1	0.26
	GSR														
	EMO											1	0.24		
	EEG					1	0.08							1	**0.19**
HAHV	ECG	1	0.30					1	0.24			2	0.32	2	0.31
	GSR														
	EMO			1	0.19					1	0.17				
	EEG					1	0.09					1	**0.12**	1	**0.14**
LAHV	ECG	1	0.23	1	0.28	1	0.29			3	**0.41**	2	0.29	2	0.36
	GSR														
	EMO													1	0.14
	EEG													1	**0.17**
LALV	ECG	1	0.32	1	0.24					2	0.31	1	0.28		
	GSR														
	EMO													2	0.31
	EEG					1	0.12							1	**0.16**
HALV	ECG	1	0.33			2	**0.44**	1	0.23	1	0.28	2	0.33	1	0.26
	GSR														
	EMO	1	0.14			1	0.14					1	**0.26**	1	0.20
	EEG			1	0.10	1	**0.09**							1	**0.15**

the significantly correlating components as predictors of the dependent emotion/personality variable, and the squared correlations (R^2) of these models are also tabulated.

Examining Table 8, the relatively few (maximum of 3) number of significant predictors can be attributed to the sparse number of principal components employed for analysis. Considering correlations with A and V, more correlates are observed for A than for V overall. At least one significant correlate is noted for all modalities except GSR. ECG is found to correlate most with A, with one correlate observed for all video types. ECG also has the most number of correlates with V (one significant correlate for LAHV and LALV clips). One EMO correlate is noted for both A and V respectively in the HAHV and HALV quadrants. A solitary EEG correlate is noted for V considering HALV clips.

A larger number of physiological correlates are observed for personality traits as compared to emotional attributes. Across all five video types, the least number of correlates are noted for Agreeableness, while most correlates are noted for Openness. The ECG modality again corresponds the maximum number of correlates, while no correlates are observed for GSR. EEG and EMO correlates are mainly noted for the Openness trait. In general, a larger number of physiological correlates are noted for emotionally similar videos for all traits. Also, linear models with a significant R^2 statistic are mainly obtained with emotion-wise similar clips, suggesting that physiology–based linear models can better predict personality traits while examining user responses under similar affective conditions. Most number of significant models are obtained for Openness, while not even one significant model is obtained for Agreeableness. Finally, focusing on the significant quadrant-specific models, the best models are noted for Extraversion (0.44 with ECG features and HALV videos) and Conscientiousness (0.41 with ECG for LAHV clips). This implies that linear physiological models acquire sufficient power to moderately explain personality

variations under such conditions.

6.2 Non-linear correlates

To examine non-linear physiological correlates of emotion and personality, we performed a mutual information analysis as previously between extracted features from the four modalities and the said attributes. Given the varying number of features for each modality, we segregated the NMI distribution over all features and the emotion/personality rating using 10-bin histograms. Fig. 7 presents the first moment or the mean of the NMI histogram distribution computed over the different video sets for each emotional/personality attribute.

It is easy to note from Fig. 7 that personality attributes share more MI with the user physiological responses than A and V, similar to the linear analyses. GSR features share maximum MI with A (highest value of 0.73 for LAHV clips), while EMO features share the most MI with V (peak of 0.75 for HALV clips). In contrast, peak MI of 0.81 is noted between ECG features and *Ex*. For both emotion and personality attributes, at least one of the NMIs observed with quadrant-based videos is higher than the NMI with all movie clips, implying that a fine-grained examination of the relationship between sub-conscious physiological responses and conscious self-ratings is more informative. Focusing on affective attributes, higher MI between ratings and physiological responses is noted for A for all modalities except EMO. Among the four modalities, ECG and EMO respectively share the most and least MI with A, while EMO and EEG share the highest and least MI with V.

Focusing on the big-five personality traits, the highest NMI histogram means over all modalities are observed for *Ex* and *Con* followed by *ES*, *Agree* and *O*. This trend is strikingly similar to the pattern of MI between affective ratings and personality scores obtained in Fig. 4.3. Examining sensing modalities, ECG features share the highest
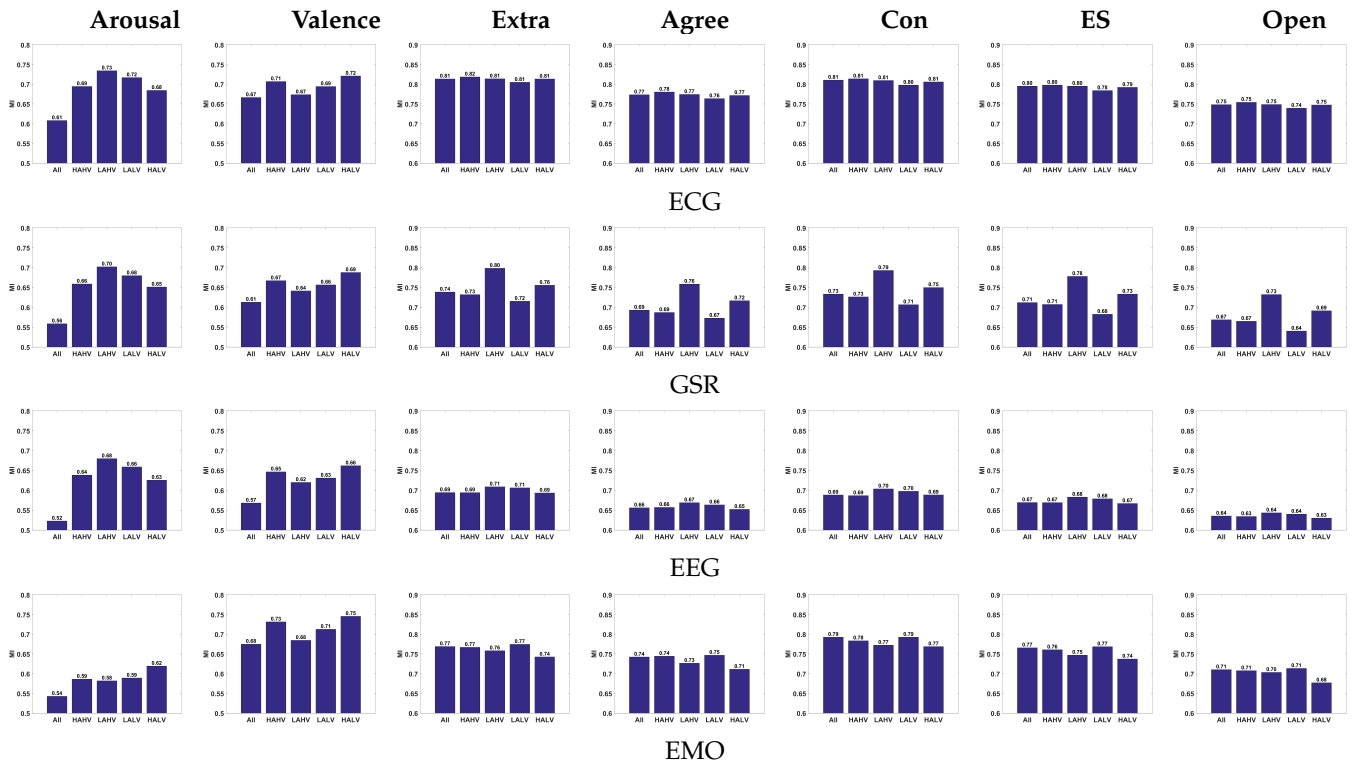
Fig. 7. (From top to bottom) Bar plots showing the means of the NMI histograms for the four modalities. Best viewed under zoom.

MI with all the personality dimensions, while EEG features correspond to the lowest NMI means.

7 RECOGNITION RESULTS

We performed binary recognition of both emotional and personality attributes to evaluate if the proposed user-centric framework can effectively achieve both. This section details the experiments and results thereof.

7.1 Emotion recognition

A salient aspect of our work is the exclusive use of commercial sensors for examining users' physiological behavior. To evaluate if our emotion recognition results are comparable to prior affective works which used laboratory-grade sensors, we followed a procedure identical to the DEAP study [10]. In particular, the most discriminative physiological features were first identified for each modality using Fisher's linear discriminant with a threshold of 0.3. Features corresponding to each user were then fed to the naive Bayes (NB) and linear SVM classifiers as shown in Table 9. A leave-one-out cross-validation scheme employed where one video is held out for testing, while the other videos are used for training. The best mis-classification cost parameter C for linear SVM is determined via grid search over $[10^{-3}, 10^3]$ again using leave-one-out cross-validation.

Table 9 presents the mean F1-scores over all users obtained using the NB and SVM classifiers with unimodal features and the decision fusion (W_{est}^t) technique described in [54]. In decision fusion, the test sample label is computed as $\sum_{i=1}^{4} \alpha_i^* t_i p_i$. Here, i indexes the four modalities used in this work, p_i's denote posterior SVM probabilities, $\{\alpha_i^*\}$ are the optimal weights maximizing the F1-score on the training

set and $t_i = \alpha_i F_i / \sum_{i=1}^{4} \alpha_i F_i$, where F_i denotes the F1-score obtained on the training set with the ith modality. Note from Section 3 that there is an equal distribution of high/low A and V, implying a class ratio (and consequently, a baseline F1-score) of 0.5

Observing Table 9, above-chance emotion recognition is evidently achieved with physiological features extracted using commercial sensors. The obtained F1-scores are superior to DEAP [10], which can possibly be attributed to (1) the use of movie clips, which are found to be better than music videos for emotional inducement as discussed in [9], and (2) to the considerably larger number of subjects employed in this study, which results in a larger training set. GSR features produce the best recognition performance for both A and V, while ECG features produce the worst recognition performance. Considering individual modalities, EEG features are better for recognizing A as compared to V, while the remaining three achieve better recognition of V. These results are consistent with earlier observations made in [9], [54]. Considering multimodal results, peripheral (ECG+GSR) features perform better than unimodal features for A recognition, while the best multimodal F1-score of 0.71 is obtained for V. Finally, comparing the two employed classifiers, NB achieves better recognition than linear SVM for both A and V.

7.2 Personality recognition

For binary personality trait recognition, we first dichotomized the big-five personality trait scores based on the median as in Section 5. This resulted in an even distribution of *high* and *low* trait labels for *Ex* and *ES*, while an inexact split for the other traits. As baselines, we consider majority-based voting and random voting according to class ratio.

Based on majority voting, baseline F1-score for the *Ex* and *ES* traits is 0.33, and 0.34 for *Ag*, 0.35 for *Con* and 0.36 for *O*. Via class-ratio based voting, a baseline score of 0.5 is achieved for all traits. We performed PCA on each feature modality in an identical fashion to linear correlation analyses prior to classification. A leave one-subject-out cross-validation scheme was used to compute the recognition results. Three classifiers were employed for recognition, i) naive Bayes, ii) linear (Lin) SVM and iii) Radial Basis Function (RBF) SVM. The C (linear and RBF SVM) and γ (RBF SVM) parameters were tuned via leave-one-subject-out grid search cross-validation on the training set.

Table 10 presents the recognition results, with the best F1-scores achieved using unimodal and multimodal features respectively denoted in bold and bold italics. For each personality trait and video set, a better-than-chance recognition F1-score (> 0.5) is achieved with at least with one of the considered modalities. Considering user physiological responses to all affective videos, the highest and lowest F1-scores are respectively achieved for *ES* (0.73) and *O* (0.53) traits– note from Fig. 2(c) that ES has the second-highest variance among the five personality dimensions, while O corresponds to the lowest variance in personality scores. Excepting for the *ES* trait, higher recognition scores are generally achieved considering user responses to emotionally similar videos, in line with the findings from linear and non-linear correlation analyses.

For all personality traits except *O*, an F1-score higher than 0.6 is achieved for at least some of the video quadrants. Among feature modalities, ECG features produce the best recognition performance across personality traits and video sets, followed by EEG, GSR and EMO. EEG features are found to be optimal for recognizing *Ex*, while ECG features achieve good recognition for the *Ag*, *Con* and *ES* traits. EMO and GSR modalities work best for the Openness trait. Focusing on classifiers, RBF SVM produces the best recognition performance for 13 out of 25 (5 personality traits × 5 video sets) conditions, while linear SVM performs best only for three conditions. Linear classifiers NB and Lin SVM perform best for the *Ex* trait, while RBF SVM, performs best for the *O* trait.

Fusion-based recognition is beneficial, and higher recognition scores are generally achieved via multimodal fusion. With user responses acquired for all videos, the highest and least fusion-based F1 scores are achieved for the *ES* (0.77 with RBF SVM) and *O* (0.56 with NB) traits respectively. With quadrant-based videos, a maximum F1-score of 0.78 is noted for *Con* (with linear SVM). NB classifier works best with fusion-based recognition, and produces best performance for the *Ex* trait achieving optimal recognition for all the five video sets.

8 DISCUSSION

The correlation analyses and recognition results clearly convey two aspects related to personality recognition from physiological data (i) A fine-grained analysis of users' physiological responses to emotionally similar movie clips enables better characterization of personality differences– this reflects in the better linear models obtained for personality traits considering quadrant-specific videos in Table 8,

and the generally higher NMIs for the same in Fig. 7. Furthermore, higher F1-scores are typically obtained when physiological responses to emotionally similar clips are used for personality trait recognition. (ii) The relationship between personality scales and physiological features is better captured via non-linear metrics– considerably high MI is noted between emotional ratings and personality scores as well as between affective physiological responses and personality traits, and this observation is reinforced with RBF-SVM producing the best recognition performance.

Interesting similarities are also evident from the correlation and recognition experiments. The NB and lin SVM classifiers work best for the Ex and ES personality traits, for which a number of linear correlates can be noted in Table 8. Also, minimum number of linear physiological correlates are noted for the *Ag* trait, for which linear classifiers do not work well (best recognition is achieved with RBF SVM for all video types except 'All' in Table 9). Likewise, no GSR correlate of emotion is observed in Table 8, which reflects in poor emotion recognition of personality traits with linear classifiers using GSR features in Table 9. Also, only some EMO correlates of personality traits are revealed in Table 8, and this modality achieves inferior personality recognition with linear classifiers.

Comparing Tables 7 and Table 9, EEG shares the least MI with all personality traits among the considered modalities, and RBF SVM performs poorly with EEG features (only one best F1 score in 25 conditions). Conversely, GSR shares considerable MI with personality dimensions, and GSR features work best with the RBF SVM classifier in Table 9. Some discrepancies also arise between the correlation and recognition results. For example, among the big-five personality traits, Openness shares the least MI with all feature modalities but has a number of linear physiological correlates. However, optimal recognition for this trait is achieved with RBF SVM, even though the achieved unimodal F1-scores are the lowest for this trait.

It is pertinent to point out some limitations of this study in general. Weak linear correlations are noted between emotional and personality scores in Table 6, and only few physiological correlates of emotion and personality are observed in Table 8, which can partly be attributed to the low variance for three of the personality dimensions and particularly the Openness trait, as seen in Fig. 2(c). In this context, median-based dichotomization of the personality scores for binary recognition may not be the most appropriate. However, most user-centered affective studies have also demonstrated recognition in a similar manner and on data compiled from small user populations, due to the inherent difficulty in conducting large-scale affective experiments. Overall, the general consistency in the nature of results observed from the correlation and recognition experiments suggest that data artifacts may have only minimally influenced our analyses, and that reliable affect and personality recognition is achievable via the extracted physiological features. Furthermore, we will make the compiled data publicly available for facilitating related research.

Even though not analyzed in this work, the ASCERTAIN database also includes Familiarity and Liking ratings, which could be useful for other research studies. For example, studying the individual and combined in-

TABLE 9
Affective state recognition with linear SVM and Naive Bayes (NB) classifiers. Mean F1-scores over all participants for the four modalities, peripheral Signals (ECG + GSR) and late fusion (W^t_{est}) are shown. Baseline F1-score is 0.5. Maximum unimodal F1-scores are shown in bold.

	ECG		GSR		EMO		EEG		Peripheral		W^t_{est}		Class Ratio
	SVM	NB	SVM	NB	SVM	NB	SVM	NB	SVM	NB	SVM	NB	
Valence	0.56	0.60	0.64	**0.68**	**0.68**	**0.68**	0.56	0.60	0.64	0.60	0.69	0.71	0.50
Arousal	0.57	0.59	0.61	**0.66**	0.59	0.61	0.58	0.61	0.62	0.69	0.64	0.67	0.50

TABLE 10
Personality recognition considering affective responses to a) all, and b) emotionally homogeneous stimuli. Maximum F1-scores with unimodal classifiers are shown in bold. Maximum fusion scores are denoted in bold italics.

Videos	Method	Extravert NB	SVM (lin)	SVM (rbf)	Agreeable NB	SVM (lin)	SVM (rbf)	Conscient NB	SVM (lin)	SVM (rbf)	Em. Stab NB	SVM (lin)	SVM (rbf)	Open NB	SVM (lin)	SVM (rbf)
All	ECG	0.56	0.06	0.53	**0.55**	0.45	0.32	**0.60**	0.51	0.55	0.53	0.60	0.58	0.48	0.35	0.49
	EEG	**0.63**	0.52	0.48	0.52	0.12	0.54	0.35	0.35	0.31	0.26	0.46	0.51	0.34	0.36	0.37
	EMO	0.35	0.31	0.45	0.40	0.35	0.42	0.39	0.36	0.34	0.44	0.36	0.47	0.50	0.36	0.26
	GSR	0.45	0.00	0.35	0.39	0.34	0.27	0.57	0.35	0.54	0.49	0.56	**0.73**	0.28	0.36	**0.53**
	W^t_{est}	*0.65*	0.52	0.57	*0.58*	0.46	0.53	0.65	0.59	*0.67*	0.59	0.64	*0.77*	*0.56*	0.42	0.53
HAHV	ECG	0.59	0.00	0.56	0.48	0.29	0.55	0.50	0.32	0.52	0.55	0.46	**0.60**	0.45	0.34	**0.55**
	EEG	**0.63**	0.43	0.63	0.54	0.10	0.10	0.34	0.35	0.33	0.19	0.32	0.57	0.41	0.35	0.45
	EMO	0.39	0.00	0.54	0.54	0.34	**0.62**	0.35	0.37	0.33	0.46	0.35	0.34	0.46	0.36	0.35
	GSR	0.22	0.00	0.31	0.47	0.34	0.51	**0.53**	0.35	0.50	0.42	0.51	0.46	0.28	0.36	0.35
	W^t_{est}	*0.65*	0.43	0.63	0.53	0.47	*0.61*	*0.60*	0.56	0.53	*0.62*	*0.62*	*0.62*	*0.59*	0.46	0.54
LAHV	ECG	0.55	0.02	0.53	0.58	0.45	**0.60**	0.70	**0.78**	0.74	**0.55**	0.46	0.41	0.56	0.42	0.49
	EEG	**0.63**	0.53	**0.63**	0.49	0.12	0.42	0.34	0.35	0.30	0.32	**0.55**	0.54	0.34	0.36	0.27
	EMO	0.49	0.34	0.51	0.43	0.35	0.10	0.58	0.37	0.36	0.51	0.35	0.39	0.46	0.37	**0.57**
	GSR	0.45	0.00	0.36	0.51	0.34	0.34	0.59	0.35	0.62	0.54	0.52	0.52	0.28	0.36	0.36
	W^t_{est}	*0.67*	0.52	0.66	*0.62*	0.49	0.60	0.74	*0.78*	0.76	*0.62*	0.60	0.61	*0.67*	0.49	0.59
LALV	ECG	0.58	0.10	0.49	0.43	0.29	0.36	0.55	0.55	**0.74**	0.53	0.58	0.50	**0.55**	0.36	0.43
	EEG	0.61	**0.63**	0.57	0.19	0.11	0.50	0.37	0.35	0.59	0.33	0.39	0.42	0.41	0.36	0.29
	EMO	0.56	0.00	0.33	0.54	0.18	**0.61**	0.30	0.37	0.36	0.49	0.35	0.49	0.33	0.36	0.48
	GSR	0.47	0.00	0.47	0.50	0.34	0.51	0.32	0.35	0.50	0.52	0.59	**0.69**	0.28	0.36	0.56
	W^t_{est}	*0.64*	0.61	0.58	0.57	0.34	*0.59*	0.58	0.56	*0.76*	0.69	0.69	*0.75*	0.57	0.46	*0.63*
HALV	ECG	0.50	0.00	0.51	0.51	0.32	**0.62**	0.57	0.57	**0.62**	0.59	0.56	**0.66**	0.45	0.33	0.50
	EEG	**0.65**	0.53	0.50	0.42	0.07	0.14	0.42	0.35	0.33	0.27	0.32	0.42	0.34	0.36	0.53
	EMO	0.32	0.34	0.30	0.47	0.35	0.47	0.42	0.36	0.40	0.33	0.36	0.44	0.56	0.36	**0.60**
	GSR	0.38	0.26	0.35	0.33	0.34	0.25	0.55	0.30	0.48	0.49	0.47	0.45	0.30	0.36	**0.60**
	W^t_{est}	*0.67*	0.57	0.59	*0.55*	0.46	*0.55*	*0.62*	0.58	0.59	0.59	0.65	*0.67*	0.57	0.42	*0.66*

fluence of familiarity, liking and personality traits on affective behavior could be relevant and useful information for recommender systems. In particular, personality-aware recommender systems have become more popular and appreciated of late [55], but the fact that personality differences show up even as consumers watch affective video content can enable video recommender systems to effectively learn user profiles over time.

Familiarity and Liking ratings could be also used to replicate and extend related studies. For example, the study presented in [56] notes a connection between familiarity, liking and the amount of smiling while listening to music. Also, Hamlen and Shuell [57] find a positive correlation between liking and familiarity for classical music excerpts, which increases when an associated video is accompanied by audio. Similar effects could be tested with emotional movie clips via ASCERTAIN.

Finally, the importance of using less-intrusive sensors in affective studies has been widely acknowledged [23], [25]. Minimally invasive and wearable sensors enable naturalistic user response, alleviating stress caused by cumbersome clinical/lab-grade equipment. Choosing minimally invasive sensors is especially critical when complex behavioral phenomena such as emotions are the subject of investigation. While most available affective datasets have been compiled using lab equipment [23], ASCERTAIN represents one of the first initiatives to exclusively employ wearable sensors for data collection, which not only enhances its ecological validity, but also repeatability and suitability for large-scale user profiling.

9 CONCLUSION

We present ASCERTAIN– a new multimodal affective database comprising implicit physiological responses of 58 users collected via commercial and wearable EEG, ECG, GSR sensors, and a webcam while viewing emotional movie clips. Users' explicit affective ratings and big-five personality trait scores are also made available to examine the impact of personality differences on AR. Among AR datasets, AS-CERTAIN is the first to facilitate study of the relationships among physiological, emotional and personality attributes.

The personality–affect relationship is found to be better characterized via non-linear statistics. Consistent results are obtained when physiological features are employed for analyses in lieu of affective ratings. Finally, AR performance superior to prior works employing lab-grade sensors is achieved (possibly because of the larger sample size used in this study), and above-chance personality trait recognition is obtained with all considered modalities. Personality differences are better characterized by analyzing responses to emotionally similar clips, as noted from both correlation and recognition experiments. Finally, RBF SVM achieves best personality trait recognition, further corroborating a non-linear emotion–personality relationship.

We believe that ASCERTAIN will facilitate future AR studies, and spur further examination of the personality–affect relationship. The fact that personality differences are observable from user responses to emotion-wise similar stimuli paves the way for simultaneous emotion and personality profiling. As recent research has shown that AR is also influenced by demographics such as *age* and *gender* [58], we will investigate correlates between affective physiological responses and the aforementioned **soft-biometrics** in future, coupled with a deeper examination on the relationship between personality and affect. We will also investigate how *a-priori* knowledge of personality can impact the design of user-centered affective studies.

10 ACKNOWLEDGEMENT

REFERENCES

[1] H. J. Eysenck, *Dimensions of Personality*, ser. The Int'l library of psychology. Transaction Publishers, 1947.

[2] W. Ng, "Clarifying the relation between Neuroticism and positive emotions," *Personality and Individual Differences*, vol. 47, no. 1, pp. 69–72, 2009.

[3] E. G. Kehoe, J. M. Toomey, J. H. Balsters, and A. L. W. Bokde, "Personality modulates the effects of emotional arousal and valence on brain activation," *Social Cognitive & Affective Neuroscience*, vol. 7, pp. 858–70, 2012.

[4] G. Stenberg, "Personality and the EEG: Arousal and emotional arousability," *Personality and Individual Differences*, vol. 13, pp. 1097–1113, 1992.

[5] V. De Pascalis, E. Strippoli, P. Riccardi, and F. Vergari, "Personality, event-related potential (ERP) and heart rate (HR) in emotional word processing," *Personality and Individual Differences*, vol. 36, pp. 873–891, 2004.

[6] A. Vinciarelli and G. Mohammadi, "A survey of personality computing," *IEEE Trans. Affective Computing*, 2014.

[7] J. Wache, R. Subramanian, M. K. Abadi, R. L. Vieriu, S. Winkler, and N. Sebe, "Implicit Personality Profiling Based on Psycho–Physiological Responses to Emotional Videos," in *Int'l Conference on Multimodal Interaction*, 2015.

[8] J. J. Gross and R. W. Levenson, "Emotion elicitation using films," *Cognition & Emotion*, vol. 9, no. 1, pp. 87–108, 1995.

[9] M. K. Abadi, R. Subramanian, S. Kia, P. Avesani, I. Patras, and N. Sebe, "DECAF: MEG-based Multimodal Database for Decoding Affective Physiological Responses," *IEEE Trans. Affective Computing*, vol. 6, no. 3, 2015.

[10] S. Koelstra, C. Mühl, M. Soleymani, J.-S. Lee, A. Yazdan, T. Ebrahimi, T. Pun, A. Nijholt, and I. Patras, "DEAP: A Database for Emotion Analysis Using Physiological Signals," *IEEE Trans. Affective Computing*, vol. 3, no. 1, pp. 18–31, 2012.

[11] M. Soleymani, J. Lichtenauer, T. Pun, and M. Pantic, "A multimodal database for affect recognition and implicit tagging," *IEEE Trans. Affective Computing*, vol. 3, pp. 42–55, 2012.

[12] M. K. Abadi, J. A. M. Correa, J. Wache, H. Yang, I. Patras, and N. Sebe, "Inference of Personality Traits and Affect Schedule by Analysis of Spontaneous Reactions to Affective Videos," in *Face and Gesture*, 2015.

[13] E. Douglas-Cowie, R. Cowie, I. Sneddon, C. Cox, O. Lowry, M. Mcrorie, J.-C. Martin, L. Devillers, S. Abrilian, A. Batliner, N. Amir, and K. Karpouzis, "The HUMAINE Database: Addressing the Collection and Annotation of Naturalistic and Induced Emotional Data," in *Affective Computing and Intelligent Interaction*, 2007, pp. 488–500.

[14] A. Batliner, C. Hacker, S. Steidl, E. Nöth, S. D'Arcy, M. J. Russell, and M. Wong, "You stupid tin box -children interacting with the AIBO robot: A cross-linguistic emotional speech corpus." in *Language Resources and Evaluation*, 2004.

[15] A. Batliner, K. Fischer, R. Huber, J. Spilker, and E. Nöth, "How to find trouble in communication," *Speech communication*, vol. 40, no. 1, pp. 117–143, 2003.

[16] C. M. Lee and S. S. Narayanan, "Toward detecting emotions in spoken dialogs," *Speech and Audio Processing*, vol. 13, no. 2, pp. 293–303, 2005.

[17] P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar, and I. Matthews, "The extended Cohn-Kanade dataset (CK+): A complete dataset for action unit and emotion-specified expression," in *CVPR Workshops*, 2010, pp. 94–101.

[18] L. Yin, X. Wei, Y. Sun, J. Wang, and M. J. Rosato, "A 3D facial expression database for facial behavior research," in *Face and Gesture*, 2006, pp. 211–216.

[19] M. Pantic, M. Valstar, R. Rademaker, and L. Maat, "Web-based database for facial expression analysis," in *Int'l Conference on Multimedia and Expo*, 2005.

[20] G. Caridakis, L. Malatesta, L. Kessous, N. Amir, A. Raouzaiou, and K. Karpouzis, "Modeling naturalistic affective states via facial and vocal expressions recognition," in *Int'l Conference on Multimodal Interaction*, 2006, pp. 146–154.

[21] N. Fragopanagos and J. G. Taylor, "Emotion recognition in human–computer interaction," *Neural Networks*, vol. 18, no. 4, pp. 389–405, 2005.

[22] N. Sebe, I. Cohen, T. Gevers, and T. S. Huang, "Emotion recognition based on joint visual and audio cues," in *Int'l Conference on Pattern Recognition*, vol. 1, 2006, pp. 1136–1139.

[23] S. Wang and Q. Ji, "Video affective content analysis: a survey of state-of-the-art methods," *IEEE Trans. on Affective Computing*, vol. 6, no. 4, 2015.

[24] J. Kim and E. Andre, "Emotion recognition based on physiological changes in music listening," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 30, no. 12, pp. 2067–2083, 2008.

[25] C. L. Lisetti and F. Nasoz, "Using noninvasive wearable computers to recognize human emotions from physiological signals," *EURASIP Journal on Advances in Signal Processing*, vol. 2004, no. 11, pp. 1672–1687, 2004.

[26] H. R.-Tavakoli, A. Atyabi, A. Rantanen, S. J. Laukka, S. Nefti-Meziani, and J. Heikkil, "Predicting the valence of a scene from observers eye movements," *PLoS ONE*, vol. 10, no. 9, pp. 1–19, 2015.

[27] R. Subramanian, D. Shankar, N. Sebe, and D. Melcher, "Emotion modulates eye movement patterns and subsequent memory for the gist and details of movie scenes," *Journal of Vision*, vol. 14, no. 3, 2014.

[28] P. T. J. Costa and R. R. McCrae, *NEO-PI-R professional manual: Revised NEO personality and NEO Five-Factor Inventory (NEO-FFI)*. Odessa, Florida: Psychological Assessment Resources, 1992, vol. 4.

[29] S. Argamon, S. Dhawle, M. Koppel, and Pennbaker, "Lexical predictors of personality type," in *Interface and the Classification Society of North America*, 2005.

[30] D. Olguin, B. Waber, T. Kim, A. Mohan, K. Ara, and A. Pentland, "Sensible organizations: Technology and methodology for automatically measuring organizational behavior," *IEEE Trans. Systems, Man, and Cybernetics-Part B: Cybernetics*, vol. 39, no. 1, pp. 43–55, 2009.

[31] X. Alameda-Pineda, J. Staiano, R. Subramanian, L. Batrinca, E. Ricci, B. Lepri, O. Lanz, and N. Sebe, "SALSA: A novel dataset for multimodal group behavior analysis," *IEEE Trans. Pattern Analysis and Machine Intelligence*, 2016.

[32] B. Lepri, R. Subramanian, K. Kalimeri, J. Staiano, F. Pianesi, and N. Sebe, "Connecting meeting behavior withExtraversion - A systematic study," *IEEE Trans. Affective Computing*, vol. 3, no. 4, pp. 443–455, 2012.

[33] R. Subramanian, Y. Yan, J. Staiano, Lanz, Oswald, and N. Sebe, "On the relationship between head pose, social attention and personality prediction for unstructured and dynamic group interactions," in *Int'l Conference on Multimodal Interaction*, 2013, pp. 3–10.

[34] G. Zen, B. Lepri, E. Ricci, and O. Lanz, "Space speaks: towards socially and personality aware visual surveillance," in *ACM Int'l Workshop on Multimodal Pervasive Video Analysis*, 2010, pp. 37–42.

[35] F. Mairesse, M. A. Walker, M. R. Mehl, and M. R. K., "Using linguistic cues for the automatic recognition of personality in conversation and text," *Journal of Artificial Intelligence Research*, vol. 30, pp. 457–500, 2007.

[36] J. Staiano, B. Lepri, N. Aharony, F. Pianesi, N. Sebe, and A. Pentland, "Friends don't lie: Inferring personality traits from social

network structure," in *ACM Conference on Ubiquitous Computing*, 2012, pp. 321–330.

[37] R. Srivastava, J. Feng, S. Roy, S. Yan, and T. Sim, "Don't ask me what i'm like, just watch and listen," in *ACM Int'l Conference on Multimedia*, 2012, pp. 329–338.

[38] A.-M. Brouwer, M. G. V. Schaik, J. E. H. Korteling, J. B. F. V. Erp, and A. Toet, "Neuroticism, Extraversion, Conscientiousness and Stress: Physiological Correlates," *IEEE Trans. Affective Computing*, vol. 6, no. 2, pp. 109–117, 2015.

[39] K. A. Winter and N. A. Kuiper, "Individual differences in the experience of emotions," *Clinical Psychology Review*, vol. 17, no. 7, pp. 791–821, 1997.

[40] S. Tok, M. Koyuncu, S. Dural, and F. Catikkas, "Evaluation of International Affective Picture System (IAPS) ratings in an athlete population and its relations to personality," *Personality and Individual Differences*, vol. 49, no. 5, pp. 461–466, 2010.

[41] S. Gupta and J. Nicholson, "Simple visual reaction time , personality strength of the nervous system : theory approach," *Personality and Individual Differences*, vol. 6, no. 4, pp. 461–469, 1985.

[42] C. C. Brumbaugh, R. Kothuri, C. Marci, C. Siefert, and D. D. Pfaff, "Physiological correlates of the Big 5: Autonomic responses to video presentations," *Applied Psychophysiology and Biofeedback*, vol. 38, no. 4, pp. 293–301, 2013.

[43] T. Lischetzke and M. Eid, "Why extraverts are happier than introverts: The role of mood regulation," *Journal of personality*, vol. 74, no. 4, pp. 1127–1162, 2006.

[44] G. Kraaykamp and K. v. Eijck, "Personality, media preferences, and cultural participation," *Personality and Individual Differences*, vol. 38, no. 7, pp. 1675–1688, 2005.

[45] J. W. Shim and B. Paul, "Effects of personality types on the use of television genre," *Journal of Broadcasting & Electronic Media*, vol. 51, no. 2, pp. 287–304, 2007.

[46] M. Perugini and L. Di Blas, "Analyzing personality-related adjectives from an eticemic perspective: the big five marker scale (BFMS) and the Italian AB5C taxonomy," *Big Five Assessment*, pp. 281–304, 2002.

[47] H. Joho, J. Staiano, N. Sebe, and J. M. Jose, "Looking at the viewer: analysing facial activity to detect personal highlights of multimedia contents," *Multimedia Tools and Applications*, vol. 51, no. 2, pp. 505–523, 2011.

[48] Y. Benjamini and Y. Hochberg, "Controlling the false discovery rate: a practical and powerful approach to multiple testing," *Journal of the Royal Statistical Society Series B (Methodological)*, vol. 57, no. 1, pp. 289–300, 1995.

[49] R. F. Bornstein, "Exposure and affect: Overview and meta-analysis of research, 1968–1987." *Psychological bulletin*, vol. 106, no. 2, p. 265, 1989.

[50] A. Aluja, O. Garca, and L. F. Garca, "Relationships among Extraversion, Openness to experience, and sensation seeking," *Personality and Individual Differences*, vol. 35, no. 3, pp. 671–680, 2003.

[51] C. Studholme, D. Hill, and D. Hawkes, "An overlap invariant entropy measure of 3D medical image alignment," *Pattern Recognition*, vol. 32, no. 1, pp. 71 – 86, 1999.

[52] P. T. Costa and R. R. McCrae, "Influence of Extraversion and Neuroticism on Subjective Well-Being: Happy and Unhappy People." *Journal of Personality and Social Psychology*, vol. 38, no. 4, p. 668, 1980.

[53] I. Mervielde, F. De Fruyt, and S. Jarmuz, "Linking openness and intellect in childhood and adulthood," *Parental descriptions of child personality: Developmental antecedents of the Big Five*, pp. 105–126, 1998.

[54] S. Koelstra and I. Patras, "Fusion of facial expressions and EEG for implicit affective tagging," *Image and Vision Computing*, vol. 31, no. 2, pp. 164–174, 2013.

[55] R. Hu and P. Pu, "A Comparative User Study on Rating vs. Personality Quiz based Preference Elicitation Methods," in *Intelligent User Interfaces*, 2009, pp. 367–371.

[56] C. V. Witvliet and S. R. Vrana, "Play it again Sam: Repeated exposure to emotionally evocative music polarises liking and smiling responses, and influences other affective reports, facial emg, and heart rate," *Cognition and Emotion*, vol. 21, no. 1, pp. 3–25, 2007.

[57] K. R. Hamlen and T. J. Shuell, "The effects of familiarity and audiovisual stimuli on preference for classical music," *Bulletin of the Council for Research in Music Education*, pp. 21–34, 2006.

[58] S. Sullivan, A. Campbell, S. B. Hutton, and T. Ruffman, "What's good for the goose is not good for the gander: age and gender differences in scanning emotion faces," *Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, pp. 1–6, 2015.

Ramanathan Subramanian received his Ph.D. degree in Electrical and Computer engineering from the National University of Singapore in 2008. He is an Associate Professor at the International Institute of Information Technology, Hyderabad (India), and previously served as a Research Scientist at UIUC's Advanced Digital Sciences Center, Singapore. His research focuses on Human-centered and Human-assisted computing. Specifically, he is interested in developing applications which utilize implicit human annotations such as eye-gaze and brain-based signatures for media and user analytics. He is a Senior Member of IEEE and an ACM member.

Julia Wache received her B.Sc. in Biology from Freie Universitat Berlin in 2009, and her M.Sc. in Biology with focus on Cognitive Science at the Humboldt University of Berlin, Germany, in 2012, working on number cognition and embodiment. She received her PhD in ICT at the University of Trento 2016 working on identifying affective states from neurophysiological signals when consuming multimedia content. Especially, she is interested in the role that personality plays in understanding and predicting human emotions.

Mojtaba Khomami Abadi is a PhD candidate at the Department of Information Engineering and Computer Science, University of Trento, Italy. Mojtaba is also the CTO of Sensaura Inc., a Canadian startup on real-time and multimodal emotion recognition technologies. His research interests include: user centric affective computing in human computer interaction and affective multimedia analysis.

Radu L. Vieriu received his B.Sc., M.Sc. and PhD degrees from "Gheorghe Asachi Technical University of Iasi, Romania, in 2009, 2010 and 2012 respectively. He is currently a postdoctoral researcher at DISI, University of Trento, Italy. His research interests span aspects from Artificial Intelligence with a focus on Machine Learning for Human Behavior Understanding.

Stefan Winkler is Distinguished Scientist and Director of the Video & Analytics Program at the University of Illinois Advanced Digital Sciences Center (ADSC) in Singapore. Prior to that, he co-founded a start-up, worked for a Silicon Valley company, and held faculty positions at the National University of Singapore and the University of Lausanne, Switzerland. He has published over 100 papers and the book Digital Video Quality (Wiley). He is an Associate Editor of the IEEE Transactions on Image Processing.

Nicu Sebe is a full professor in the University of Trento, Italy, where he is leading the research in the areas of multimedia information retrieval and human behavior understanding. He was a General Co-Chair of FG 2008 and ACM MM 2013, and a program chair of CIVR 2007 and 2010, and ACM MM 2007 and 2011. He is a program chair of ECCV 2016 and ICCV 2017. He is a senior member of IEEE and ACM and a fellow of IAPR.