

Assessment of Personality Dimensions Across Situations in Dyadic Role-Play Scenarios

Alice Zhang, *Student Member, IEEE*, Skanda Muralidhar, Daniel Gatica-Perez, *Member, IEEE*, and Mathew Magimai-Doss, *Member, IEEE*

Abstract—Research indicates that users prefer assistive technologies whose personalities align with their own. This has sparked interest in automatic personality perception (APP), which aims to predict an individual's perceived personality traits. Prior studies in APP have treated personalities as static traits, independent of context. However, perceived personalities can vary by situation as shown in psychological research. In this study, we investigate the relationship between conversational speech and perceived personality for participants engaged in two work situations (a neutral interview and a stressful client interaction). Our key findings are: 1) perceived personalities differ significantly across interactions, 2) loudness, sound level, and spectral flux features are indicative of perceived Extraversion, Agreeableness, Conscientiousness, and Openness in neutral interactions, while Neuroticism correlates with these features in stressful contexts, 3) handcrafted acoustic features and non-verbal features outperform speaker embeddings in inference of perceived personality, and 4) stressful interactions are more predictive of Neuroticism, aligning with existing psychological research.

Index Terms—apparent personality perception, computational paralinguistics

I. INTRODUCTION

MODELING human personality is fundamental to the development of affective computing systems capable of personalized interactions. Users are more engaged with and have greater trust in assistive technologies that reflect their own personalities, thereby demonstrating the value of personality-aware systems [1], [2]. Consequently, there is a growing field of research on automatic personality perception (APP), the task of inferring one's personality as perceived by external judges. Unlike automatic personality recognition (APR), which focuses on inferring self-reported personality, APP captures perceived personality and better reflects the cues that affective computing systems are designed to interpret.

Describing an individual's personality is a complex task. The Five-Factor Model of Personality (Big-5) is a widely accepted description of personality traits in five dimensions: Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness to Experience [3]. A number of previous works have leveraged this description of personality to infer traits using modalities such as speech [4], [5], facial expressions [6], body language [7], and physiological signals [8]. The APP task is not new, as there exists a body of personality

computing works [9]; however, existing works have mainly utilized datasets in which subjects are annotated in a single context. For instance, in the widely used Speaker Personality Corpus [4], [10], individuals are rated solely based on their behaviors in news broadcasts. Thus, analyses from these datasets assume that personalities are static and independent of external factors.

On the other hand, psychology and organizational behavior researchers have long recognized that personality expression can be situation dependent [11], [12], [13]. Experts in these fields have spent decades researching and debating the stability of traits across roles and environments as seen in the personality-situation debate. While researchers have since generally acknowledged an interactionist view in which both an individual's traits and the given situation influence expression of personality through behavior, affective computing models often lack mechanisms to adjust for context [12], [13].

Despite the importance of studying personality in relation to situational context, computational analyses of personality across situations remain limited. Addressing this gap, we utilize the UbImpressed dataset in which participants engaged in both job-related neutral and stressful conversations and received personality ratings for both interactions [14]. Given the conversational nature of this dataset, we aim to pose the question: *To what extent do conversational features (e.g., speech, non-verbal cues) explain perceived personality dimensions across varying contexts?* To answer this, we investigate the following research questions (RQ):

RQ1 Is there a significant difference in ratings of perceived personality of the same participants across two job-related conversation scenarios?

RQ2 How does the relationship between perceived personality and conversational features vary across contexts?

RQ3 How do conversational features from different interactions differ in their inference of perceived personality?

To our knowledge, while other researchers have studied constructs such as job performance on the two situations in this dataset, ours is the first to investigate perceived personality across two different situations in this dataset [15], [16].

The paper is structured as follows: section II discusses relevant work in affective computing and psychology, section III describes the dataset, section IV discusses the methodology for three analyses performed to study the research questions, section V presents results of each analysis, and section VII concludes the paper.

A. Zhang is with Idiap Research Institute, Switzerland and The University of Texas at Austin, USA. (email: alice.zhang@austin.utexas.edu)

D. Gatica-Perez is with Idiap Research Institute and EPFL, Switzerland.

S. Muralidhar and M. Magimai-Doss are with Idiap Research Institute, Switzerland.

II. RELATED WORK

A. Contextual Variability in Expression of Personality Traits

Researchers in psychology have long discussed the extent to which personality traits versus situational factors play a role in shaping behavior as demonstrated by the person-situation debate [11]. Proponents of the “person” perspective argue that individuals exhibit relatively consistent behaviors across situations, suggesting that underlying traits drive these behavioral patterns. In contrast, advocates of the “situation” perspective found that individuals’ behaviors are relatively inconsistent across time and situation, arguing that personality traits do not exist and behaviors are influenced by the situation more than the individual’s disposition [17].

While many resolutions have been proposed, a synthesis resolution acknowledging that both personality traits and situational factors influence expressed behavior is generally accepted [12], [13], [17]. Over long periods of time, traits predict behavior well and explain differences in behavior between people. In momentary behavior, an individual’s behavior is variable and traits may not strongly describe behaviors [18]. Thus, both perspectives are necessary for a full understanding of personality. However, existing works in APP have mainly focused on personality as a static trait while psychology research has shown that situational factors also influence expressed personality. For more comprehensive computing solutions to infer personality, the situation in which the solution is employed also needs to be taken into consideration.

B. Explaining Contextual Variability in Personality Within Job-Related Settings

Two theories have been proposed to explain variations in personality-related responses across job-relevant contexts: the situation strength principle [19] and trait activation theory (TAT) [20]. The situation strength principle proposes that strong situations, defined as environments in which rules, structures or cues provide clear guidance on an individual’s expected behavior, restrict expression of personality. Meanwhile, weak situations, in which there are fewer cues regarding expected behavior, allow for greater expression of personality [19], [21]. Complementing this theory, TAT states that individuals express specific traits when situations offer opportunities for a specific trait to be expressed. For instance, an individual with high Extraversion may exhibit more extroverted behavior when working in a sales role involving customer interaction.

Complementary to research around dimensions of personality, recent psychological research has given rise to frameworks, such as DIAMONDS [22] and CAPTIONS [23], that characterize situations along a set of dimensions (e.g., *complexity* - does the situation require deep thinking?). Furthermore, dimensions of situations predict which personality traits are most readily displayed. For instance, pleasant situations most strongly allow for the manifestation of behaviors consistent with Extraversion, Agreeability, and Openness, thereby suggesting that individuals engage in more prosocial behaviors in positive situations [22], [23].

This body of psychological work illustrates the significance of studying personality within its situational context. While prior investigations of the UbImpressed dataset used perceived

personality to predict hirability and job performance [14], [24], our work aims to predict personality from conversational behaviors and examine how this prediction task is influenced by the context of the interaction.

C. Automatic Personality Perception using Speech

As speech signals encode rich information in addition to the spoken content itself, speech is a promising modality for the APP task. For instance, the 2012 INTERSPEECH Speaker Trait Challenge [10] resulted in a body of different approaches towards personality inference using speech ranging from low-level acoustic descriptors to spectrum analysis [25], [26]. More recent works showed that short utterance filler words [27] and dictionary learning of spectrograms [28] can be used to classify speaker’s personality traits. However, prior works treat personalities as a static trait and independent of context. They utilize datasets in which only one set of personality ratings was collected per participant, whether the data came from crowd-sourced monologue interview responses [5], clips of video blogs from YouTube [29], [30], or speech from news clips [4], [10], [25], [26].

In contrast to these static approaches, recent work in emotion recognition has demonstrated the importance of contextual information [31], [32]. However, this context-aware paradigm has not yet been fully explored in personality computing. While isolated studies have examined the relationship between personality and speech tasks (e.g., reading neutral texts versus commenting on thematic apperception test images [33]), there is an absence of work that compares the same individuals across varying contexts to quantify the shift in conversational behaviors for personality computing. Our study fills this gap by utilizing a paired-scenario design to isolate the effect of context on personality perception.

III. DATASET

We use the UbImpressed dataset [14], a corpus of dyadic interactions designed to study workplace behavior. This is the only dataset to our knowledge to contain ratings for the same participants across distinct conversational contexts, allowing for intra-subject analysis of perceived personality. While newer large-scale datasets exist, they neither allow for intra-subject analysis [34] nor contain unscripted speech [34], which are essential to our research questions. Thus, despite its age, UbImpressed is uniquely suited for this investigation.

A. Participants and Protocol

The study cohort consists of 100 students (57 female, 43 male) from a hospitality management school, with a mean age of 20.6 years (SD = 2.47). Reflecting the school’s bilingual curriculum, the dataset contains conversations in English (23%) and French (77%). The data collection was part of a two-part behavioral training program. All 100 participants completed the first session, and 69 participants completed the second session, for a total of 169 sessions. In each session, students engaged in two role-play scenarios:

- 1) **Employment interview:** The student acted as an applicant for a hospitality internship, while a research assistant played the recruiter. Questions focused on

TABLE I
INTRA-CLASS CORRELATION FOR PERCEIVED PERSONALITIES AND STRESS.

Session	Scenario	ICC(2,k)					
		Extra	Agree	Consc	Neuro	Open	Stress
1	Interview	0.65	0.54	0.59	0.51	0.48	0.74
	Desk	0.62	0.72	0.69	0.56	0.33	0.63
2	Interview	0.66	0.50	0.56	0.45	0.37	0.54
	Desk	0.77	0.58	0.59	0.37	0.58	0.45

motivation and experience. While job interviews can be stressful, this one was framed as a practice opportunity.

- 2) **Hotel reception desk interaction with an unsatisfied customer:** The student acted as a hotel receptionist handling a client (played by a research assistant), who was unsatisfied with charges on their bill and hostile. This interaction was confrontational and more stressful compared to the *interview*. The increased stress of this interaction is empirically verified in our analysis of personality ratings (section V-A).

Following the first session, students received feedback from professionals in human resources (HR) on their performance. The dataset contains a total of 338 interactions (169 sessions \times 2 scenarios). Although the scenarios were simulated and the interlocutor’s speech was scripted, the participants’ speech was entirely spontaneous. The average duration of the *interview* was 7.6 minutes and 4.6 minutes for the *desk*. From here on, we refer to *scenario* as the *interview* or *desk* interaction and the *session* as the first (1) or second (2) instance of an interaction.

B. Perceived Personality Rating

Perceived personality was manually annotated based on the Big-5 traits and stress. Annotators were Master’s students in the same psychology program, ensuring a shared understanding of the trait definitions. The rating process was performed using 2-minute slices of the audiovisual recording, a common practice in psychology [35]. To minimize bias between scenarios, the *interview* and *desk* interactions were annotated by independent groups of annotators (five raters for *interviews*; three for *desk* interactions). No annotator rated the same participant in both scenarios. Each participant received a score on a Likert-scale of 1 (Low) to 7 (High) for each dimension. As shown in Table I, inter-rater reliability (assessed via Intraclass Correlation Coefficient, ICC(2, k)) indicated at least moderate Agreement ($ICC \geq 0.5$) for most traits.

IV. METHODS AND EXPERIMENTAL SETUP

A. Personality Rating Comparison across Conversations

To understand whether perceived personality differs across conversations (RQ1), we analyze the distribution of ratings across the *interview* and *desk* scenarios and the first and second sessions of each interaction. We employ the two-sample Kolmogorov-Smirnov (KS) test for goodness of fit where the null hypothesis states that the underlying continuous distributions $F(x)$ and $G(x)$ of two independent samples are identical for all x [36]. We first use the two-sample KS test to compare the distributions of participants’ perceived stress across different scenarios and sessions to verify that the two interaction scenarios in the UbImpressed dataset are significantly different to validate the use of the UbImpressed dataset

for our research questions. Then, we use the two-sample KS test to compare the distribution of perceived personality across different scenarios and sessions to answer RQ1.

B. Feature Extraction and Selection

We diarize each conversation and keep only the audio segments corresponding to speech from the students. From each student’s speech, we extract features that have been shown to vary with a speaker’s emotional state and influenced by personality traits [37]. While recent advancements in speech processing heavily utilize deep learning embeddings to maximize accuracy, these representations lack interpretability. Because our objective is to identify how specific behavioral mechanisms drive personality perception under varying conversational contexts, we prioritize interpretable feature sets. Consequently, we select (1) eGeMAPS features and (2) non-verbal features. To contextualize the performance of these handcrafted features against newer representations, we also benchmark them against (3) speaker embeddings. The specific dimensions corresponding to each category are available here¹.

eGeMAPS [38] features are a set of acoustic features commonly used for speech-based affective computing using the openSMILE toolkit [39]. It contains 88 features that capture frequency, energy, amplitude, and spectral parameters initially hand-crafted for speech emotion recognition.

Non-verbal features include 75 audio and visual cues selected for their relevance in existing literature in psychology [40] and social computing [41], and previously extracted for the UbImpressed dataset. We utilize this set because it provides a multimodal behavioral profile of a speaker that eGeMAPS features alone cannot address. The feature categories include: speaking activity (21 dimensions; e.g. turn duration, pause statistics), prosody (30 dimensions; e.g., spectral entropy, voicing statistics), head nods (8 dimensions), visual back-channeling (6 dimensions) and overall visual motion (10 dimensions). These features fundamentally capture conversation-level dynamics and are thus extracted at the conversation-level.

Speaker embeddings are fixed-dimensional representations of speech that encode speaker identity. Work by Ulgen et al. [42] show that intra-speaker clusters of ECAPA-TDNN vectors [43], a type of speaker embedding, can capture affective states. Thus, we use 512-dimensional ECAPA-TDNN vectors extracted via the Pyannote toolkit [44].

We extract eGeMAPS and speaker embedding features at the utterance (turn) level, following prior literature [38], [43]. To model the overall interaction, we then aggregated these turn-level features into a single, conversation-level vector by calculating the mean or median of each feature dimension across a participant’s conversation. We found that personality inference using the median feature vector consistently outperformed the mean aggregation. Thus, our primary analyses utilize this median vector to represent the interaction.

To investigate the temporal dynamics of personality expression and capture how acoustic markers evolve over the course of an interaction, we conducted an additional temporal

¹https://osf.io/8vztd/?view_only=57e6b2ae1602457a89bf9472889a4d64

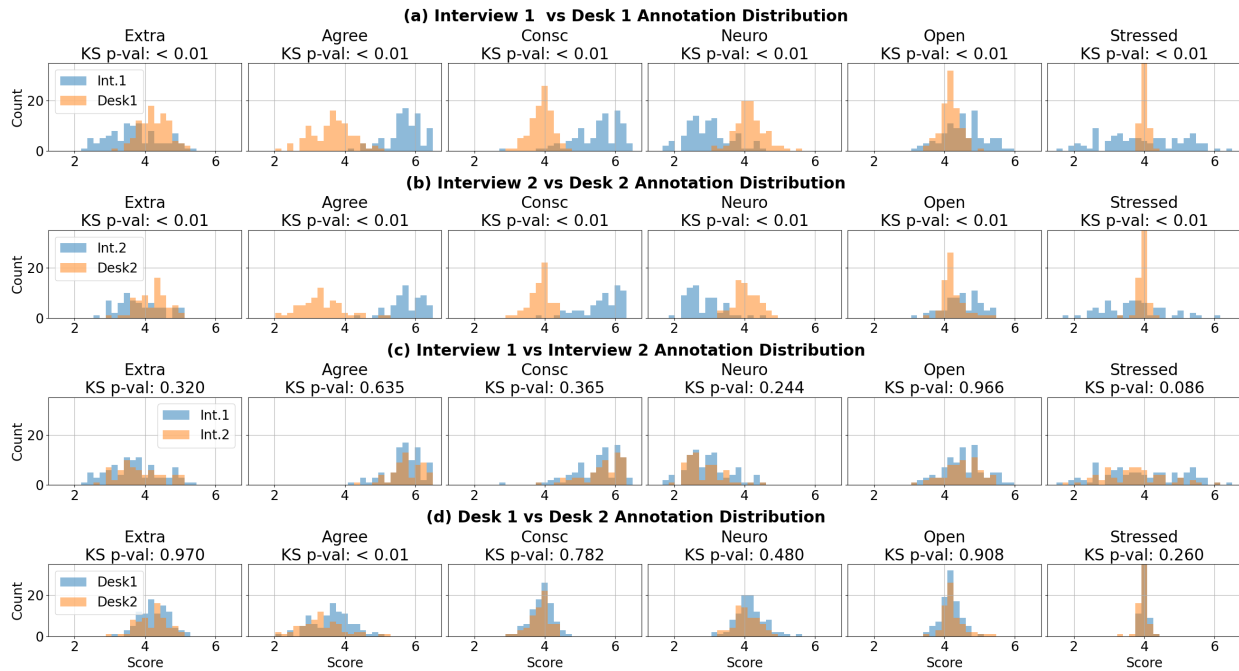


Fig. 1. (RQ1) Comparing distribution of perceived personality scores between: (a) interview and desk scenario of first lab session, (b) interview and desk scenario of second lab session, (c) interview scenario of first and second lab sessions, and (d) desk scenario of first and second lab sessions.

analysis with the eGeMAPS features, since they generally outperformed speaker embeddings. We isolated feature vectors from five distinct temporal landmarks based on the chronological sequence of each participant’s turns: the first utterance, the 25th percentile utterance, the midpoint (50th percentile utterance), the 75th percentile utterance, and the last utterance.

On the other hand, the non-verbal feature vector, which contains audio and video-based features, captures conversation-level features of the participant’s interaction. Therefore, there is only one feature vector extracted for each participant per interaction, and we use the feature vector as is.

Lastly, for each feature set, we select a subset of features using the Pearson Correlation Coefficient (PCC). The PCC measures the strength of the linear relationship between two continuous variables. We assume that relevant information is contained in features significantly correlated with perceived personality scores and select features with $p < 0.05$ (RQ2).

C. Experimental Setup for Inference of Perceived Personality

We formulate the inference of perceived personality from speech signals as a regression task (RQ3) to compare personality inference across conversations.

1) *Conversation-Specific Inference*: First, we aim to infer perceived personality on a specific conversation scenario and session to understand how feature selection impacts performance. We perform feature selection as detailed in section IV-B using data from the complementary session of the same scenario so that no data from the test set influences the feature selection process. For instance, to build and evaluate a system inferring personality in the first *interview* session, we perform feature selection using data from the second *interview* session.

To infer a continuous score for each personality dimension, we train regression models using one of the three feature sets. Our primary architecture is a Random Forest (RF) Regressor

with 200 trees, depth of 5, and 16 samples per leaf node. We selected a RF because it captures non-linear relationships and provides interpretable feature importance metrics. To ensure the robustness of our model selection, we benchmarked the RF against ElasticNet, K-Nearest Neighbors (KNN) and Support Vector Machine regressors and conducted hyperparameter optimization (e.g., tuning maximum depth for the RF) of all models to ensure fair comparison. The RF consistently provided competitive or superior predictive performance across the Big-5 traits, empirically justifying its use as the primary model for our downstream feature importance analysis.

We evaluate all regressors using the coefficient of determination R^2 , which is the amount of total variance explained by the model, and the PCC between the predicted and observed personality scores in a 10-fold cross validation scheme. Each feature vector corresponds to one participant, so each evaluation fold consists of speakers unseen in the training process. Thus, the evaluation is speaker independent. We obtain confidence intervals by bootstrapping the evaluation set.

2) *Inference across Scenarios and Sessions*: To understand the generalization of personality regression across different scenarios within the same set of participants, we build a RF regressor trained on data from one scenario (e.g. *interview 1*) and evaluate using data from the complementary scenario of the same session (e.g. *desk 1*). We perform feature selection using the same method as in Sections IV-B and IV-C1.

Then, to understand how a trained RF regressor generalizes across conversation instances of the same scenario over time, we build a regressor trained on data from one session (e.g. *interview 1*) and evaluate it using data from the complementary session of the same scenario (e.g. *interview 2*). We perform feature selection using the training dataset. In this setup, the training and testing sets are not strictly participant-

independent, as the 69 participants who returned for the second session appear in both sets. However, because each session was annotated separately and participants modified their behavior based on HR feedback, both behavior and target labels differ across sessions. This setup evaluates the model's capacity to generalize to behavioral shifts within a cohort (see Figure 1 in Supplementary Material).

V. RESULTS AND ANALYSIS

A. Personality Rating Comparison Results and Analysis

As shown in Figure 1, we reject the null hypothesis of the two-sample KS test in favor of the alternative hypothesis when comparing all personality dimensions and stress scores across *interview* and *desk* scenarios. In contrast, we fail to reject the null hypothesis in comparing personality and stress across the first and second sessions of the *interview* or *desk* scenario, with the exception of Agreement in the *desk* scenario. This analysis shows that participants are perceived to exhibit significantly more stress in the client interaction ($\mu = 4.06$) compared to the interview ($\mu = 3.82$) and validates the suitability of the UbImpressed dataset to study our research questions.

Overall, this analysis shows that perceived personality differs significantly between the neutral interview and stressful client interaction for the same speaker (RQ1). In the *interview*, participants were perceived as more agreeable ($\mu = 5.81$ vs. 3.53) and open ($\mu = 4.52$ vs. 4.19), and less neurotic ($\mu = 2.92$ vs. 4.11). While psychological frameworks have established that personality expression varies by situation [22], [23], these studies rely on participants' retrospective self-reports of past events. Our results provide empirical validation that observers perceive these personality shifts from interactions.

B. Correlation Analysis Results and Analysis

TABLE II
(RQ2) CORRELATION OF SELECT eGEMAPS AND NON-VERBAL FEATURES WITH PERSONALITY. $\dagger p < 0.01$; $* p < 0.05$.

Feature	Scenario	Extra	Agree	Consc	Neuro	Open
<i>equivalent Sound Level (dB)</i> (eGeMAPS)	Int. 1	0.52 \dagger	0.43 \dagger	0.31 \dagger	-0.08	0.41 \dagger
	Int. 2	0.41 \dagger	0.34 $*$	0.37 \dagger	-0.27 $*$	0.27 $*$
	Desk 1	0.09	0.41 \dagger	0.04	-0.36 \dagger	0.19
	Desk 2	0.21	0.49 \dagger	0.15	-0.27 $*$	0.07
# pauses >1 second (non-verbal)	Int. 1	0.18	0.23 $*$	0.16	-0.07	-0.27 $*$
	Int. 2	0.31 $*$	0.22	0.04	-0.13	0.38 \dagger
	Desk 1	-0.12	0.23 $*$	-0.03	-0.27 $*$	0.24 $*$
	Desk 2	0.07	0.15	0.07	-0.02	0.02
Count Head Nods (non-verbal)	Int. 1	0.42 \dagger	0.25 $*$	0.24 $*$	-0.06	0.31 \dagger
	Int. 2	0.44 \dagger	0.01	-0.07	-0.07	0.29 \dagger
	Desk 1	0.25 $*$	0.07	0.10	-0.24 $*$	0.08
	Desk 2	0.39 \dagger	0.22	0.04	-0.30 $*$	0.09

Select results of the correlation between each participant's median eGeMAPS and non-verbal feature vector, and personality dimensions are presented in Table II (RQ2). The complete correlation matrix for all feature sets can be found here ¹.

For both sessions, there are stronger correlations between eGeMAPS and nonverbal features and personality in the *interview* for all personality dimensions except Neuroticism, in which the correlation between Neuroticism and features is stronger in the *desk* scenario. Additionally, there is an inverse in correlation direction between Neuroticism and features. We hypothesize this is due to the association between Neuroticism

and negative emotions such as anxiety, in which prior research has shown that anxiety introduces irregularities in conversational behaviors compared to positive or neutral speech [45].

For eGeMAPS features, there is a common subset of features that are significantly correlated with all personality dimensions except Neuroticism in the *interview*. Energy- (loudness and equivalent sound level) and spectral- (spectral flux) related features are all correlated with perceived Extraversion, Agreement, Conscientiousness, and Openness in the *interview*. In contrast, the correlation is significant between this subset of features and Neuroticism only in the *desk* scenario. Interestingly, significant correlations between this subset of speech features and Agreement exist in both the *interview* and *desk* interaction, suggesting that this trait may be more consistently expressed across contexts.

For non-verbal features, we observe that features related to speech energy, voicing rate, and head nods are positively correlated with all personality dimensions except Neuroticism in the *interview*. In the *desk* scenario, we observe positive correlations between features related to speaking activity, number of pauses, and speech energy, and Agreement but negative correlations between these features and Neuroticism.

C. Inference of Perceived Personality Results

1) *Conversation-Specific Perceived Personality*: We observe that the RF regressor contains some predictive power for all personality dimensions across all three feature sets (Table III; eGeMAPs and non-verbal results only, speaker embedding results in Supplementary Material). Generally, eGeMAPS and non-verbal features outperform speaker embeddings, with non-verbal features explaining a maximum 30% of the variance for the Extraversion dimension. Our eGeMAPS results are comparable to the results obtained by Barchi *et al.* who report R^2 values between 0.12 and 0.22 for each personality dimension [29]. We prioritize this comparison because [29] isolates the performance of eGeMAPS features on spontaneous speech. To our knowledge, no comparable works on personality regression using non-verbal and speaker embedding features exist.

Our temporal analysis reveals different dynamics in the contexts for session 1 (Figure 2, see Supplementary Material for session 2). In the *interview*, the predictive performance is relatively stable across time. The median feature vector generally achieves the most robust performance across traits (notably for Extraversion, Agreeableness, and Conscientiousness in session 1). We also observe that the first and last utterances yield localized peaks in performance for Extraversion and Agreeableness in the first session. This aligns with first and last impressions often setting the tone of an interaction and thus carrying concentrated social signals. In contrast, for the *desk* scenario, the temporal dynamics were more volatile. For traits like Agreement, Neuroticism, and Openness, some utterances from temporal landmarks consistently outperformed the median aggregate. These results show that personality expression in stressful scenarios is inconsistent over time and markers of personality appear at specific moments.

Our regressor performs comparably with and without feature selection, showing that a subset of features can have as much predictive power as the entire feature set. Additionally, across

TABLE III

(RQ3) REGRESSION RESULTS FOR PREDICTING PERSONALITY DIMENSIONS WITHIN A GIVEN SCENARIO AND SESSION. CELLS WITH A DIMENSION (DIMS) OF FEATURES SET TO 0 INDICATE THERE ARE NO FEATURES WITH STATISTICALLY SIGNIFICANT CORRELATIONS WITH THE PERSONALITY TRAIT. BOLD HIGHLIGHTS THE BEST PERFORMANCE FOR THE GIVEN PERSONALITY DIMENSION ACROSS BOTH SCENARIOS. $\dagger p < 0.01$; $* p < 0.05$.

Feature	Session	Extraversion			Agreement			Conscientiousness			Neuroticism			Openness		
		Dims	R^2	r	Dims	R^2	r	Dims	R^2	r	Dims	R^2	r	Dims	R^2	r
Hand-crafted (eGeMAPS)	1	88	0.25 ± 0.14*	0.51 [†]	88	0.16 ± 0.13*	0.41 [†]	88	0.07 ± 0.10, 0.28	88	-0.04 ± 0.10, 0.07	88	0.16 ± 0.12, 0.39 [†]			
		27	0.27 ± 0.14*	0.52 [†]	52	0.19 ± 0.14*	0.42 [†]	39	0.01 ± 0.13, 0.19	15	-0.05 ± 0.12, 0.05	18	0.19 ± 0.15*			
	2	88	0.02 ± 0.17, 0.22	88	0.18 ± 0.12*	0.45 [†]	88	0.24 ± 0.13, 0.51*	88	-0.02 ± 0.10, 0.11	88	-0.08 ± 0.15, 0.03				
		31	0.03 ± 0.17, 0.25*	43	0.20 ± 0.12*	0.47[†]	28	0.18 ± 0.13, 0.47 [†]	6	-0.08 ± 0.11, 0.01	32	-0.06 ± 0.16, 0.09				
Non-verbal	1	75	0.29 ± 0.12, 0.55 [†]	75	0.18 ± 0.13*	0.41 [†]	75	0.07 ± 0.11, 0.27 [†]	75	-0.05 ± 0.09, 0.02	75	0.22 ± 0.12, 0.46[†]				
		35	0.30 ± 0.12*	0.55[†]	23	0.19 ± 0.13*	0.43 [†]	15	0.06 ± 0.12, 0.26 [†]	2	-0.05 ± 0.09, -0.01	18	0.20 ± 0.14, 0.44 [†]			
	2	75	0.17 ± 0.11*, 0.55 [†]	75	0.07 ± 0.11, 0.44 [†]	75	0.10 ± 0.15, 0.34 [†]	75	-0.03 ± 0.09, 0.08	75	0.04 ± 0.11, 0.24*					
		32	0.12 ± 0.15*, 0.38 [†]	35	0.08 ± 0.11, 0.32 [†]	28	0.10 ± 0.17, 0.35 [†]	6	-0.07 ± 0.08, -0.08	36	0.06 ± 0.12, 0.28*					

(a) Interview

Feature Type	Session	Extraversion			Agreement			Conscientiousness			Neuroticism			Openness		
		Dims	R^2	r	Dims	R^2	r	Dims	R^2	r	Dims	R^2	r	Dims	R^2	r
Hand-crafted (eGeMAPS)	1	88	-0.04 ± 0.08, 0.02	88	0.08 ± 0.11*	0.30*	88	-0.10 ± 0.07, -0.16	88	0.08 ± 0.10, 0.31 [†]	88	-0.10 ± 0.10, -0.09				
		20	-0.07 ± 0.07, -0.11	26	0.09 ± 0.12, 0.33 [†]	0	-	13	0.12 ± 0.14, 0.36[†]	1	-0.07 ± 0.07, -0.03					
	2	88	-0.03 ± 0.17, 0.13	88	0.18 ± 0.16, 0.44 [†]	88	-0.15 ± 0.08, -0.39 [†]	88	-0.07 ± 0.09, -0.01	88	-0.14 ± 0.09, -0.23					
		5	-0.09 ± 0.05, -0.38 [†]	24	0.18 ± 0.16, 0.45 [†]	2	-0.09 ± 0.05, -0.34 [†]	21	-0.01 ± 0.12, 0.16	11	-0.15 ± 0.08, -0.35 [†]					
Non-verbal	1	75	-0.01 ± 0.08, 0.09	75	0.01 ± 0.10, 0.20	75	-0.03 ± 0.08, 0.06	75	0.07 ± 0.10, 0.28 [†]	75	-0.11 ± 0.08, -0.10					
		10	0.02 ± 0.10, 0.20	16	0.01 ± 0.10, 0.23*	1	-0.09 ± 0.04, -0.27 [†]	13	0.02 ± 0.10, 0.20	2	-0.05 ± 0.08, 0.01					
	2	75	-0.04 ± 0.08, 0.01	75	-0.07 ± 0.10, -0.02	75	-0.10 ± 0.08, -0.18	75	0.02 ± 0.13, 0.22	75	-0.12 ± 0.08, -0.21					
		13	0.02 ± 0.10, 0.18	30	-0.02 ± 0.11, 0.10	3	-0.11 ± 0.05, -0.51 [†]	28	0.05 ± 0.12, 0.28*	7	-0.14 ± 0.09, -0.43 [†]					

(b) Desk

R² over time by training dataset
(Above bars = Pearson r, p < 0.05)

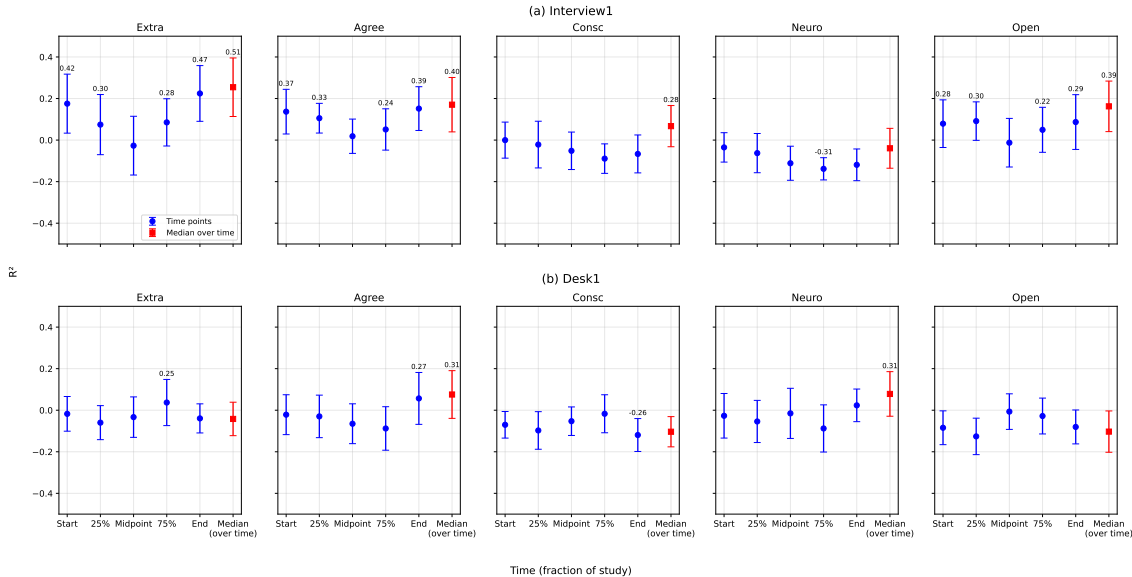


Fig. 2. Temporal dynamics of predicting personality in (a) Interview 1 and (b) Desk 1.

both sessions and all feature sets, we observe that the *desk* dataset is better for inference of Neuroticism and that the *interview* dataset is better for inference of the other four personality dimensions. These results support the correlation results presented in section V-B and also psychological research on the inference of personality from situations. Situations in which negative feelings can occur and lack enjoyment are associated with Neuroticism whereas situations containing pleasant interactions are associated with Extraversion, Agreement, and Openness [22], [23]. Additionally, stressful situations are more likely to activate the expression of Neuroticism [20].

2) Perceived Personality Across Scenarios and Sessions:

We report results of personality inference on the scenario unseen during training in Table IV (eGeMAPS only; others in Supplementary Material). The regressor performs significantly worse in this cross-scenario validation despite the training and evaluation datasets containing the same participants. For

most personality dimensions, the features do not explain any of the variance. Only non-verbal features explain a minimal amount of variance in Extraversion when training on the *interview* and evaluating on the *desk* interaction, indicating that certain non-verbal behaviors are predictive of Extraversion in both stressful and neutral interactions. But, the overall lack of generalizability emphasizes that the relationship between speech and personality varies across situations.

The results of personality inference across sessions of the same scenario are in Table V (eGeMAPS only; others in Supplementary Material). For eGeMAPS features, this evaluation setup outperforms the evaluation in which the training and test data come from the same session and scenario. The performance improvement is perhaps due to training with the entire conversation dataset rather than nine of ten folds.

We still observe that the *desk* interaction is more predictive of Neuroticism. These results combined with the results

TABLE IV

(RQ3) REGRESSION RESULTS FOR eGEMAPS. PREDICTING PERSONALITY DIMENSIONS WHEN TRAINING ON ONE SCENARIO AND EVALUATING ON THE OTHER SCENARIO WITHIN THE SAME SESSION. CELLS WITH A DIMENSION (DIMS) OF FEATURES SET TO 0 INDICATE THAT THERE ARE NO FEATURES WITH STATISTICALLY SIGNIFICANT CORRELATIONS TO THE PERSONALITY DIMENSION. † $p < 0.01$; * $p < 0.05$.

Session	Train Scenario	Eval. Scenario	Extraversion		Agreement		Conscientiousness		Neuroticism		Openness	
			Dims	R^2, r	Dims	R^2, r	Dims	R^2, r	Dims	R^2, r	Dims	R^2, r
1	Interview	Desk	88	-0.41 ± 0.29, 0.17	88	-19.1 ± 5.0, 0.11	88	-32.0 ± 8.1, 0.07	88	-8.6 ± 2.8, 0.12	88	-7.7 ± 2.0, 0.14
			27	-0.64 ± 0.34, 0.11	52	-19.9 ± 5.3, 0.13	39	-32.3 ± 7.9, -0.01	15	-8.2 ± 2.4, -0.11	18	-8.5 ± 2.3, 0.22 [†]
	Desk	Interview	88	-0.47 ± 0.26, 0.27 [†]	88	-20.5 ± 5.6, 0.31 [†]	88	-6.0 ± 2.1, 0.07	88	-4.8 ± 1.7, 0.01	88	-0.4 ± 0.22, 0.25 [*]
			20	-0.40 ± 0.28, 0.40 [†]	26	-20.5 ± 5.6, 0.22 [*]	0	-	13	-4.7 ± 1.6, 0.13	1	-0.33 ± 0.22, 0.17
2	Interview	Desk	88	-0.55 ± 0.43, 0.3 [*]	88	-20.9 ± 8.8, 0.21	88	-41.6 ± 11.5, -0.01	88	-16.1 ± 4.7, -0.07	88	-2.6 ± 1.3, -0.11
			31	-0.75 ± 0.50, 0.29 [*]	43	-20.9 ± 8.9, 0.22	28	-36.4 ± 10.0, 0.13	6	-17.0 ± 5.0, -0.04	32	-3.1 ± 1.4, 0.02
	Desk	Interview	88	-0.09 ± 0.15, 0.29 [*]	88	-35.5 ± 10.9, 0.04	88	-11.8 ± 4.5, -0.21	88	-7.33 ± 3.4, -0.03	88	-0.40 ± 0.28, 0.01
			5	-0.29 ± 0.22, 0.10	24	-35.2 ± 10.8, 0.33 [†]	2	-10.9 ± 4.1, -0.02	21	-7.3 ± 3.4, -0.08	11	-0.43 ± 0.29, 0.16

TABLE V

(RQ3) REGRESSION RESULTS FOR eGEMAPS. PREDICTING PERSONALITY DIMENSIONS WHEN TRAINING ON ONE SESSION AND EVALUATING ON THE OTHER SESSION OF THE SAME SCENARIO. † $p < 0.01$; * $p < 0.05$.

Scenario	Train Session	Eval. Session	Extraversion		Agreement		Conscientiousness		Neuroticism		Openness	
			Dims	R^2, r	Dims	R^2, r	Dims	R^2, r	Dims	R^2, r	Dims	R^2, r
Interview	1	2	88	0.09 ± 0.28 [†] , 0.41 [†]	88	0.26 ± 0.16 [†] , 0.51 [†]	88	0.18 ± 0.15 [†] , 0.53 [†]	88	-0.032 ± 0.12 [†] , 0.21	88	0.05 ± 0.17 [†] , 0.30 [*]
			31	0.10 ± 0.28 [†] , 0.41 [†]	43	0.27 ± 0.16 [†] , 0.52 [†]	28	0.10 ± 0.18 [†] , 0.43 [†]	6	-0.14 ± 0.18, 0.10	32	0.03 ± 0.18 [†] , 0.30 [*]
	2	1	88	0.17 ± 0.09 [†] , 0.56 [†]	88	0.17 ± 0.10 [†] , 0.49 [†]	88	0.07 ± 0.12 [†] , 0.34 [†]	88	-0.02 ± 0.13, 0.20 [*]	88	0.11 ± 0.06 [†] , 0.41 [†]
			27	0.19 ± 0.08 [†] , 0.57 [†]	52	0.17 ± 0.10 [†] , 0.48 [†]	39	0.07 ± 0.13, 0.34 [†]	15	-0.02 ± 0.14, 0.19	18	0.15 ± 0.08 [†] , 0.45 [†]
Desk	1	2	88	-0.04 ± 0.13, 0.14	88	-0.21 ± 0.36, 0.46 [†]	88	-0.08 ± 0.11, 0.02	88	0.01 ± 0.25 [†] , 0.29 [*]	88	-0.05 ± 0.12, 0.09
			5	-0.09 ± 0.20, 0.10	24	-0.21 ± 0.37, 0.48 [†]	2	-0.14 ± 0.18, -0.03	21	-0.07 ± 0.26 [†] , 0.29 [*]	11	-0.09 ± 0.11, 0.03
	2	1	88	-0.11 ± 0.21, 0.08	88	-0.33 ± 0.34, 0.40 [†]	88	-0.06 ± 0.09, 0.05	88	0.03 ± 0.08 [†] , 0.42 [†]	88	-0.04 ± 0.07, 0.03
			20	-0.18 ± 0.18, 0.03	26	-0.32 ± 0.34, 0.42 [†]	0	-	13	0.11 ± 0.08 [†] , 0.50 [†]	1	-0.05 ± 0.10, 0.12

from the cross-scenario evaluation highlight that, within a type of scenario, there are specific and relatively consistent relationships between speech features and perceived personality dimensions. However, the relationships do not generalize well across different conversation scenarios. In conclusion, our results demonstrate that acoustic and non-verbal markers of personality are highly context-dependent; thus, affective computing systems should move beyond static models to account for situational context and temporal dynamics of the interaction.

VI. LIMITATIONS AND FUTURE WORK

A. Ecological Validity of Role-Play

Our findings are specific to goal-oriented, dyadic interactions within a professional setting due to our use of simulated role-play interactions. However, while the interaction context was fixed, the participants' speech was unscripted and spontaneous. Furthermore, HR professionals provided performance feedback, ensuring the interactions carried meaning for the participants. Future work should investigate whether these conversational patterns persist in unrestricted, naturalistic settings.

Additionally, our analyses rely on the UbImpressed dataset and represent a specific demographic (hospitality management students). Consequently, the generalizability of our results to other populations remains to be verified. Nonetheless, our results highlight the importance of modeling context to better perceive and adapt to users. For instance, in an application for clinical interviews, distinguishing between situational anxiety and Neuroticism as a trait would be critical for diagnosis.

B. Real-World Applicability

While our modest predictive power (R^2) limits immediate real-world application, explainable variance must be evaluated against the theoretical upper bound of inter-rater agreement [46]. For example, given an Extraversion ICC of 0.65, our model ($R^2 = 0.30$) effectively captures 46% (0.30/0.65) of the explainable variance. Ultimately, our objective is not

achieving state-of-the-art accuracy, but using interpretable features to analyze personality inference shifts across contexts.

C. Speech Emotion Recognition

This study focuses on neutral and stressful contexts to isolate the effect of situational stress on personality perception. Future studies could integrate speech emotion recognition (SER) to distinguish between specific emotional states within each context. Such an analysis would disentangle whether emotional cues drive differences in conversational behavior observed in personality perception and prediction [47].

VII. CONCLUSION

We investigated the relationship between perceived personality and conversational speech using the UbImpressed dataset. Our experiments showed that perceived personality of the same participants differs significantly across a neutral and stressful interaction (RQ1). We found that features related to loudness, equivalent sound level, and spectral flux are correlated with perceived Extraversion, Agreement, Conscientiousness, and Openness in the neutral scenario (*interview*) and with perceived Neuroticism in the stressful scenario (*desk interaction*) (RQ2). Through a regression analysis, we demonstrated that eGEMAPS and non-verbal features explain up to 27% and 30% of variance respectively in perceived Extraversion. Furthermore, features better predicted Neuroticism in the *desk* scenario, but the remaining personality dimensions in the *interview* scenario (RQ3). Our results emphasize the relevance of building and evaluating APP systems relative to the context in which behaviors are observed. Furthermore, our results highlight the importance of context-aware affective computing systems, which has thus far been understudied.

ACKNOWLEDGMENT

The authors gratefully acknowledge the Swiss Federal Commission for Scholarships for funding this research with the Swiss Government Excellence Scholarship, grant 2024.0241 and Innosuisse Flagship project IICT, grant PFFS-21-47.

REFERENCES

- [1] M. Braun, A. Mainz, R. Chadowitz, B. Pflöging, and F. Alt, "At your service: Designing voice assistant personalities to improve automotive user interfaces," in *CHI '19*, 2019, p. 1–11.
- [2] E. C. Snyder, S. Mendu, S. S. Sundar, and S. Abdullah, "Busting the one-voice-fits-all myth: Effects of similarity and customization of voice-assistant personality," *Int. J. Hum. Comput. Stud.*, vol. 180, 2023.
- [3] R. R. McCrae and O. P. John, "An introduction to the five-factor model and its applications," *J. Personality*, 1992.
- [4] G. Mohammadi and A. Vinciarelli, "Automatic personality perception: Prediction of trait attribution based on prosodic features extended abstract," in *ACII 2015*, 2015, pp. 484–490.
- [5] L. Chen, R. Zhao, C. W. Leong, B. Lehman, G. Feng, and M. E. Hoque, "Automated video interview judgment on a large-sized corpus collected online," in *ACII 2017*, 2017, pp. 504–509.
- [6] A. Kachur, E. Osin, D. Davydov, K. Shutilov, and A. Novokshonov, "Assessing the big five personality traits using real-life static facial images," *Scientific Reports*, vol. 10, 2020.
- [7] R. Subramanian, Y. Yan, J. Staiano, O. Lanz, and N. Sebe, "On the relationship between head pose, social attention and personality prediction for unstructured and dynamic group interactions," in *ICMI 2013*, 2013, p. 3–10.
- [8] 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, 2018.
- [9] A. Vinciarelli and G. Mohammadi, "A survey of personality computing," *IEEE Trans. Affect. Comput.*, vol. 5, no. 3, pp. 273–291, 2014.
- [10] B. Schuller, S. Steidl, A. Batliner, E. Noeth, A. Vinciarelli, F. Burkhardt, R. van Son, F. Weninger, F. Eyben, T. Bocklet, G. Mohammadi, and B. Weiss, "The INTERSPEECH 2012 speaker trait challenge," in *Proc. Interspeech 2012*.
- [11] W. Mischel, *Personality and Assessment*. Wiley, 1968.
- [12] N. S. Endler and D. Magnusson, "Toward an interactional psychology of personality," *Psychological bulletin*, vol. 83, pp. 956–74, 1976.
- [13] W. Mischel and P. Peak, "Some facets of consistency: Replies to Epstein, Funder, and Bem," *Psychological Review*, vol. 90, pp. 394–402, 1983.
- [14] S. Muralidhar, L. S. Nguyen, D. Frauendorfer, J.-M. Odobez, M. Schmid Mast, and D. Gatica-Perez, "Training on the job: behavioral analysis of job interviews in hospitality," in *ICMI '16*, 2016, p. 84–91.
- [15] S. Muralidhar, M. S. Mast, and D. Gatica-Perez, "A tale of two interactions: Inferring performance in hospitality encounters from cross-situation social sensing," *IMWUT*, vol. 2, no. 3, Sep. 2018.
- [16] S. Muralidhar, R. Siegfried, J.-M. Odobez, and D. Gatica-Perez, "Facing employers and customers: What do gaze and expressions tell about soft skills?" in *MUM 2018*, p. 121–126.
- [17] W. Fleeson and E. Nofle, "The end of the person–situation debate: An emerging synthesis in the answer to the consistency question," *Soc. Personal. Psychol. Compass*, 2008.
- [18] W. Fleeson, "Moving personality beyond the person-situation debate: The challenge and the opportunity of within-person variability," *Curr. Dir. Psychol. Sci.*, 2004.
- [19] W. Mischel, E. B. Ebbeesen, and A. R. Zeiss, "Selective attention to the self: Situational and dispositional determinants," *J. Pers. Soc. Psychol.*, vol. 27, no. 1, p. 129, 1973.
- [20] R. Tett and D. Burnett, "A personality trait-based interactionist model of job performance," *J. Appl. Psychol.*, 2003.
- [21] T. A. Judge and C. P. Zapata, "The person–situation debate revisited: Effect of situation strength and trait activation on the validity of the big five personality traits in predicting job performance," *Academy of Management Journal*, vol. 58, no. 4, pp. 1149–1179, 2015.
- [22] J. Rauthmann, D. Gallardo-Pujol, E. Guillaume, E. Todd, C. Nave, R. Sherman, M. Ziegler, A. Jones, and D. Funder, "The situational eight DIAMONDS: A taxonomy of major dimensions of situation characteristics," *J. Pers. Soc. Psychol.*, vol. 107, pp. 677–718, 2014.
- [23] S. Parrigon, S. E. Woo, L. Tay, and T. Wang, "CAPTION-ing the situation: A lexically-derived taxonomy of psychological situation characteristics," *J. Pers. Soc. Psychol.*, vol. 112, 2017.
- [24] S. Muralidhar, M. Schmid Mast, and D. Gatica-Perez, "How may I help you? Behavior and impressions in hospitality service encounters," in *ICMI*, 2017, p. 312–320.
- [25] J. Pohjalainen, S. Kadioglu, and O. Räsänen, "Feature selection for speaker traits," in *Proc. Interspeech*, 09 2012, pp. 270–273.
- [26] A. Ivanov and X. Chen, "Modulation spectrum analysis for speaker personality trait recognition," in *Proc. Interspeech 2012*, pp. 278–281.
- [27] M. Tayarani, A. Esposito, and A. Vinciarelli, "What an "ehm" leaks about you: Mapping fillers into personality traits with quantum evolutionary feature selection algorithms," *IEEE Trans. Affect. Comput.*, vol. 13, no. 1, pp. 108–121, 2022.
- [28] M.-A. Carbonneau, E. Granger, Y. Attabi, and G. Gagnon, "Feature learning from spectrograms for assessment of personality traits," *IEEE Trans. Affect. Comput.*, vol. 11, no. 1, pp. 25–31, 2020.
- [29] R. Barchi, L. Pepino, M. Gauder, L. Estienne, P. Riera, and L. Ferrer, "Apparent personality prediction from speech using expert features and wav2vec 2.0," in *SMM 2023*, pp. 21–25.
- [30] J.-I. Biel and D. Gatica-Perez, "The YouTube lens: Crowdsourced personality impressions and audiovisual analysis of Vlogs," *IEEE Trans. on Multimedia*, vol. 15, no. 1, pp. 41–55, 2013.
- [31] R. Kosti, J. M. Alvarez, A. Recasens, and A. Lapedriza, "Emotion recognition in context," in *CVPR 2017*, 2017, pp. 1960–1968.
- [32] J. Lee, S. Kim, S. Kim, J. Park, and K. Sohn, "Context-aware emotion recognition networks," in *ICCV*, 2019.
- [33] A. Guidi, C. Gentili, E. Scilingo, and N. Vanello, "Analysis of speech features and personality traits," *Biomed. Signal Process. Control*, vol. 51, pp. 1–7, 05 2019.
- [34] L. Sun, J. Zhao, and Q. Jin, "Revealing personality traits: A new benchmark dataset for explainable personality recognition on dialogues," in *EMNLP*, Nov. 2024, pp. 19 988–20 002.
- [35] N. Ambady and R. Rosenthal, "Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis," *Psychological Bulletin*, vol. 111, no. 2, pp. 256–274, 1992.
- [36] J. L. Hodges, "The significance probability of the Smirnov two-sample test," *Arkiv för Matematik*, vol. 3, pp. 469–486, 1958.
- [37] F. Mairesse, M. A. Walker, M. R. Mehl, and R. K. Moore, "Using linguistic cues for the automatic recognition of personality in conversation and text," *J. Artif. Int. Res.*, vol. 30, no. 1, p. 457–500, Nov. 2007.
- [38] F. Eyben, K. Scherer, B. Schuller, J. Sundberg, E. Andre, C. Busso, L. Devillers, J. Epps, P. Laukka, S. Narayanan, and K. Truong, "The Geneva minimalistic acoustic parameter set (GeMAPS) for voice research and affective computing," *IEEE Trans. Affect. Comput.*, 2016.
- [39] F. Eyben, M. Wöllmer, and B. Schuller, "openSMILE – the munich versatile and fast open-source audio feature extractor," in *ACM Multimedia 2010*, pp. 1459–1462.
- [40] T. DeGroot and J. Gooty, "Can nonverbal cues be used to make meaningful personality attributions in employment interviews?" *Journal of Business and Psychology*, vol. 24, no. 2, pp. 179–192, Jun. 2009.
- [41] L. S. Nguyen, D. Frauendorfer, M. S. Mast, and D. Gatica-Perez, "Hire me: Computational inference of hirability in employment interviews based on nonverbal behavior," *IEEE Trans. on Multimedia*, vol. 16, no. 4, pp. 1018–1031, 2014.
- [42] I. Ulgen, Z. Du, C. Busso, and B. Sisman, "Revealing emotional clusters in speaker embeddings: A contrastive learning strategy for speech emotion recognition," in *ICASSP 2024*, 04 2024, pp. 12 081–12 085.
- [43] B. Desplanques, J. Thienpondt, and K. Demuynck, "ECAPA-TDNN: Emphasized channel attention, propagation and aggregation in tdnn based speaker verification," *Proc. Interspeech 2020*.
- [44] H. Bredin, R. Yin, J. M. Coria, G. Gelly, P. Korshunov, M. Lavechin, D. Fustes, H. Titeux, W. Bouaziz, and M.-P. Gill, "pyannote.audio: neural building blocks for speaker diarization," in *ICASSP 2020*.
- [45] T. Özseven, M. Düğenci, A. Doruk, and H. I. Kahraman, "Voice traces of anxiety: acoustic parameters affected by anxiety disorder," *Archives of Acoustics*, pp. 625–636, 2018.
- [46] H. Nili, C. S. Wingfield, A. Walthers, L. Su, W. D. Marslen-Wilson, and N. Kriegeskorte, "A toolbox for representational similarity analysis," *PLoS Computational Biology*, vol. 10, no. 4, p. e1003553, 2014.
- [47] K. H. Greenaway, E. K. Kalokerinos, and L. A. Williams, "Context is everything (in emotion research)," *Soc. Personal. Psychol. Compass*, vol. 12, no. 6, p. e12393, 2018.

Alice Zhang is a Ph.D. candidate at the University of Texas at Austin. Her research interests include acoustic sensing and wearable technologies to support and analyze social interactions.

Skanda Muralidhar (Ph.D. EPFL) is a researcher at Idiap Research Institute, where he researches and develops AI-powered technologies for sign language learning. His current work advances applied AI for education, healthcare, and human-centered innovation.

Daniel Gatica-Perez directs the Social Computing Group at Idiap Research Institute and is a professor at EPFL. His research integrates human-centered and participatory methods with mobile, social, and AI technologies to support individuals and communities.

Mathew Magimai-Doss (Ph.D. EPFL) is a Senior Researcher at Idiap Research Institute, Switzerland. His research interests lie in signal processing, statistical pattern recognition, artificial neural networks and computational linguistics with applications to speech, audio, and multimodal signal processing.