Signal, Image and Video Processing (2026) 20:383
https://doi.org/10.1007/s11760-026-05442-y

ORIGINAL PAPER

The impact of facial expression and head orientation on personality perception

Melike Kara[1] · Sinan Sonlu[1] · Arçin Ülkü Ergüzen[1] · Funda Durupinar[2] · Uğur Güdükbay[1]

Received: 4 February 2026 / Revised: 8 May 2026 / Accepted: 14 May 2026

Abstract

This study investigates personality expression in neural-rendered talking-head videos. We generated eight video clips with four different expressions and two different head poses. We conducted a user study comparing pairs of these videos to evaluate differences in perceived Big Five traits across various facial expressions and head orientations. Findings indicate that facial expressions have a significant impact on perceived personality, with happy expressions receiving higher scores on socially positive traits. On the other hand, head orientation had a limited overall effect on personality perception. Participants predominantly selected the "equal" option across conditions, indicating that orientation alone rarely produced big perceptual differences. However, subtle directional effects emerged. These findings provide practical insights for the design of digital humans and virtual agents, highlighting that expressive facial cues play a dominant role in shaping personality impressions, while head orientation serves as a secondary, context-dependent modulator. These findings establish a perceptual baseline for how specific parameters in neural-rendered videos influence human trait attribution.

Keywords Facial expression · Head movement · Personality perception · Talking-head synthesis · Conversational agent

1 Introduction

Human interaction is deeply influenced by rapid personality impressions, which occur in less than a second [1]. Evidence suggests that even 'thin slices' of expressive behavior-brief observations lasting less than five minutes-provide sufficient information for observers to make accurate social judgments [2]. These impressions shape the social judgments of individuals even when they lack access to verbal content or background information [3]. Such impressions are derived not only from dynamic facial expressions but also from the head orientation, which can alter the perceived meaning [4]. Gaining insights into how human impressions arise is crucial for the entertainment industry, particularly in computer animation and game design, as well as in the design of virtual agents and human-computer interaction systems based on visual communication.

Works in traditional animation and comics highlight the significance of facial expressions and head movements in shaping our perception of personalities. McCloud [5, 6] explains how subtle changes in facial expressions can influence the perception of a character's personality in comics. Thomas and Johnston [7] describe the fundamental principles of animation, emphasizing how facial expressions and head orientation contribute to the clarity of characters' personalities.

Recent advances in neural talking-head synthesis and face reenactment have enabled the generation of highly realistic animated portraits that replicate human facial movements and head dynamics [8–11]. This is achieved by separating a person's appearance from motion representations. These movements are encoded in latent spaces; the encoded features capture facial geometry, articulation, and temporal coher-

✉ Uğur Güdükbay
gudukbay@cs.bilkent.edu.tr

Melike Kara
melike.kara@ug.bilkent.edu.tr

Sinan Sonlu
sinan.sonlu@bilkent.edu.tr

Arçin Ülkü Ergüzen
ulku.erguzen@bilkent.edu.tr

Funda Durupinar
funda.durupinarbabur@umb.edu

1 Department of Computer Engineering, Bilkent University, Ankara, Türkiye

2 Department of Computer Science, University of Massachusetts, Boston, MA, USA

Published online: 02 June 2026

ence, often via keypoints or dense motion representations [8, 10]. Although new models demonstrate the importance of generating highly realistic and interactive talking heads [12], more knowledge is required on how specific nonverbal facial cues in neural-rendered faces influence perceived personality traits [11, 13]. As early motion models often struggle to correctly capture complex appearance transformations in keypoint neighborhoods [8], even subtle rendering inconsistencies in modern pipelines may unintentionally shape the observer's social impressions during large head orientation changes.

Previous studies on personality perception showed that facial appearance, expression, and movements provide systematic clues related to the Big Five personality dimensions [14–16]. Many existing studies simultaneously vary multiple modalities, such as identity, facial expression, head pose, speech content, and voice characteristics, making it difficult to isolate the contribution of individual visual cues to perceived personality [11, 17]. This confounding of factors is significant in studies involving visual, auditory, and linguistic channels that are often tightly coupled.

We focus on isolating the visual contribution of facial expression and head orientation to perceived personality in neural-rendered talking-head videos. We deliberately remove speech audio and linguistic variation, and adjust identity, background, lighting, and stimulus duration across all conditions. While it is known that speech and voice characteristics contribute to personality perception, these have been intentionally omitted from neural-rendered videos to minimize potential effects and facilitate a more controlled study. This design enables the direct attribution of perceptual differences to nonverbal facial motion cues and head orientation encoded by the neural reenactment process.

To implement this controlled approach, we utilize high-quality driving signals from the speech subset of the RAVDESS corpus [18], which provides professionally validated emotional expressions. All auditory and linguistic channels were removed to isolate the impact of nonverbal facial motion and head dynamics on personality perception, so it prevents cross-modal dominance effects [19]. Using left and front image of neutral base identity from the Multi-PIE dataset [20, 21] and the LivePortrait [10] neural reenactment model, we generate short talking-head video clips. The experimental stimuli are systematically structured across two dimensions: facial expression (angry, happy, sad, surprised) and head orientation (frontal vs. left-facing). Participants evaluate pairs of videos within a forced-choice pairwise comparison framework. They select the video that appears higher on each of the Big Five personality traits, as well as perceived naturalness, or state that each is perceived the same for related traits.

We address the following research questions and formulate the following hypotheses based on prior findings in facial expression and personality perception literature:

- **RQ1:** Does facial expression influence perceived personality traits in neural-rendered talking-head videos?
- **RQ2:** Does head orientation influence perceived personality traits when facial expression is held constant?
- **RQ3:** Which personality traits exhibit greater sensitivity to expression or head orientation variations?
- **H1:** Facial expression will significantly affect perceived personality traits, with positive expressions (e.g., happiness) associated with higher perceived extraversion and agreeableness, and negative expressions (e.g., anger) associated with lower ratings.
- **H2:** Head orientation will influence perceived personality, such that frontal-facing videos will be perceived as higher in extraversion and agreeableness than side-facing videos.
- **H3:** The magnitude of these effects will vary across personality dimensions, with extraversion and agreeableness exhibiting larger effect sizes than conscientiousness or openness.

This work empirically examines how observers interpret facial expressions and head poses in terms of personality traits in neural-rendered videos. By focusing on perception rather than synthesis quality alone, our work contributes to research at the intersection of computer graphics, social perception, and human-computer interaction, with potential implications for the design of digital humans and embodied agents.

2 Related work

Personality Perception from Visual Cues. Evidence from cognitive neuroscience suggests that the human visual system differentiates between static facial features, like identity, and dynamic signals, such as expression and head pose, through specialized neural circuits that facilitate the decoding of social intent [22]. Numerous studies indicate that facial appearance and expression contribute to systematic impressions related to the Big Five traits, remaining consistent even when observers lack biographical or linguistic context [1, 14, 17]. According to the 2D model of face evaluation, these impressions are structured around the dimensions of valence and dominance, serving as adaptive mechanisms to infer behavioral intentions and physical capabilities [3]. Nonverbal dynamics, such as facial motion and head orientation, influence perceived personality, suggesting that traits like extraversion and agreeableness may be sensitive to expressive cues [15, 23]. Even a neutral face can convey complex social and emotional messages depending on its orientation [4].

Furthermore, recent studies examine how body motion and expressive animations affect user engagement and perceptions of personality in conversational agents [24]. In interactive and embodied agent settings, personality perception is commonly shaped by multiple channels simultaneously, such as voice, dialogue style, gesture, eye gaze, and facial expression [25–27]. Recent datasets provide large-scale personality labels for automated recognition through deep learning by aggregating cues such as gaze and action units [28]. While such multimodal designs increase practical relevance, they also introduce overlap of influences that make it difficult to attribute perceived trait differences to specific visual factors [11, 17]. In contrast, our study employs a controlled, ceteris paribus experimental design to isolate the visual contributions of facial expressions and head poses. This controlled approach allows us to establish direct links between subtle cues and social impressions that might be confounded in more complex, aggregate datasets.

Neural Talking-Heads and Face Reenactment. Neural-rendered talking head approaches support identity-preserving animation and facial motion transfer from driving signals such as audio or source video [8–10]. Modern approaches typically disentangle identity-specific appearance from motion representations in a latent space. From a geometric perspective, these models often rely on unsupervised keypoint detection (with local affine transforms) and/or dense motion fields (e.g., optical-flow-like warps) to map source appearance onto target poses, linking surface deformation and temporal coherence to perceived intent and expressivity [8, 11]. Recent state-of-the-art frameworks, such as TIMAR [12], have shifted the focus to the interactive aspects of head generation by unifying speaking and listening behaviors.

Alongside model accuracy [29], the graphics community emphasizes that perceptual factors significantly influence how users interpret rendered content. Perception-driven graphics and accelerated rendering surveys highlight how human sensitivity and artifacts shape what is noticed and believed [30]. A recent survey on realistic human animation consolidates how rendering, motion realism, and evaluation protocols interact to create believable digital humans [31].

Despite these technical advances, comparatively less is understood about how specific *isolated* nonverbal cues in neural-rendered faces (e.g., expression versus head pose) shape perceived Big Five personality traits. Most prior systems vary multiple factors simultaneously, making it difficult to determine whether observed trait differences arise from facial motion cues themselves or from correlated changes in identity, audio/linguistic content, or synthesis artifacts [11, 17]. Our work, on the other hand, separately evaluates the effects of expression and head orientation, while controlling for identity, background, lighting, duration, and audio. We additionally incorporate perceived naturalness as a control measure to decouple personality effects from rendering quality [11, 32].

3 Stimuli generation

To isolate the effects of *facial expression* and *head pose* on perceived personality, while controlling for identity, speech content, background, and duration, we generated a tightly controlled set of neural-rendered talking-head videos using a fixed identity and carefully selected driving motions.

Base Identity. Throughout the study, we used one identity represented by two portrait photographs of the same person, captured from frontal and left views. The images feature a neutral facial expression under uniform lighting conditions and only include the person's head. We also adjusted the facial structure, hairstyle, and background to match the same identity. The selected identity was a single male and displayed a neutral facial expression under uniform lighting conditions. Only the head region was included.

Identity images were selected from the CMU Multi-PIE face dataset [20, 21, 33], which provides controlled, high-quality face images captured under uniform lighting conditions and a consistent background. While using a single identity limits generalizability, it provides a clean experimental setting that allows us to directly attribute perceptual changes to facial expression and head pose, the primary focus of this work.

Driving Videos and Emotion Selection. To obtain realistic facial motion and head dynamics, we selected short driving segments from a controlled-emotional speech subset of the RAVDESS corpus, which provides professionally acted facial expressions under consistent recording conditions [18]. These videos capture natural speech-driven facial motion and spontaneous head movement under unconstrained conditions, making them suitable as motion sources for neural reenactment. Although the original driving video clips contained neutral sentences, we selected clips without the actor's voice to eliminate the effect of the actor's emotional tone in our experiments.

The labels in the RAVDESS corpus define the intended emotional expressions. However, we did not perform a separate perceptual validation of the reenacted outputs; small deviations in perceived expression introduced by the reenactment process cannot be completely ruled out.

We consider four emotional categories that are commonly studied in facial expression research: *anger*, *happiness*, *sadness*, and *surprise*. For each emotion, we manually selected driving clips that clearly and dominantly express that emotion. While these expressions are professionally posed, such behaviors are universally recognizable and effectively communicate intended emotional states [34]. Posed expressions often represent high-intensity versions of spontaneous emotions, ensuring clear signal salience in experimental settings [34]. All selected segments were chosen to exhibit comparable temporal duration and expressive salience, thereby minimizing systematic intensity differences across conditions.

Neural Reenactment. In our reenactment setup, the source identity is provided by two neutral facial images of the same identity from the CMU Multi-PIE dataset, and the driving signal is provided by a short emotional video segment selected from the RAVDESS corpus. Each selected driving video is combined with the base identity images using the LivePortrait neural reenactment model [10]. LivePortrait disentangles identity-specific appearance features from motion features and synthesizes a talking-head video in which the reference identity follows the driver's facial expression and head orientation dynamics. Following the LivePortrait framework, identity-specific appearance information is extracted from the static source image. The motion-related information is derived from the driving video via implicit keypoints that encode changes in facial expression and head pose. The model then transfers these motion patterns to the source identity through neural warping and decoding, while preserving the source person's facial appearance.

LivePortrait's stitching and retargeting components help improve spatial consistency and fine control over eye and lip motion [10]. From a geometric standpoint, the reenactment process implicitly models facial surface deformation using dense motion fields and keypoint-based correspondence, enabling manipulation of expression and head orientation while preserving identity-specific shape. The final output of this process creates a short, synthesized talking-head video. LivePortrait's dense motion representation effectively captures high-fidelity expressions while maintaining identity-preserving realism. Therefore, we ensure that the perceived personality remains driven by intended social cues rather than by rendering artifacts such as smoothing effects, which can lead to a loss of subtle nuances and potentially trigger the Uncanny Valley effect [35, 36].

Figure 1 depicts the neural reenactment pipeline used in this study. All videos have the same resolution and frame rate. Their durations range between three and four seconds. We muted the clips to ensure that only visual cues influenced the participants. During the experiment, videos were looped continuously to allow participants to observe motion patterns without time pressure.

Pair Construction. From the eight base conditions, we constructed ten comparison pairs designed to probe two complementary questions: the effect of head orientation within the same emotion, and the effect of facial expression under a fixed head pose. Specifically, we define:

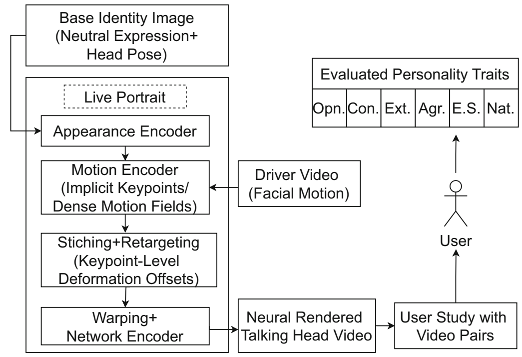

Fig. 1 Overview of the neural reenactment pipeline based on LivePortrait. Identity-specific appearance features are extracted from a neutral base image, while facial expression and head orientation are encoded from a driving video using implicit keypoints and dense motion fields. Stitching and retargeting modules apply keypoint-level deformation offsets before neural warping and decoding, producing a rendered talking-head video with preserved identity and controlled motion

- *Same-emotion, different head orientation pairs (4):* frontal versus left-facing video clips for each emotion (e.g., *happy-front* vs. *happy-left*), isolating the effect of head orientation while holding expression constant.
- *Different-emotion, same head orientation pairs (6):* all unordered pairs of emotions at frontal view (e.g., *angry-front* vs. *happy-front*), isolating the effect of facial expression while holding head orientation constant.

4 Experiments

We conducted a user study to collect data on how facial expressions and head orientations of neural-rendered talking head videos influence perceived personality traits in humans. The study used a forced-choice, pairwise-comparison paradigm. The appendix includes screenshots from the user study and the questions asked during the study.

Participants and Procedure. We recruited N=80 English-fluent participants via Prolific. Of the 77 who reported demographics, 39 were female and 38 male, with a mean age of 45.47±13.63 years; 55 were from the UK and 22 from the USA. Desktop or laptop usage was required for eligibility. A sample size of 80 ensures sufficient statistical power for both the pairwise emotion-comparison rankings and the ordinal preference analysis of head orientation. For the former, we fitted a Bradley-Terry (BT) model with Davidson extension. The sample size provided 80 observations per unique emotional pair (480 total comparisons). This exceeds the requirement for stable parameter estimation in paired-comparison models, which typically requires 30–50 observations per pair to reliably detect a medium difference in worth parameters (Δ=0.5) while accounting for a tie parameter δ [37]. For ordinal preference analysis, we fitted a cumulative link mixed model (CLMM) and assessed power using the proportional odds framework [38]. With 80 participants providing 320 total observations, the study was powered to detect a medium effect size (odds ratio$\approx$2.0) with α=0.05 and $1-\beta>0.80$, even after accounting for the design effect due to repeated measures. Thus, the final sample was considered robust for evaluating both orientation and expressive cues.

The study was conducted through a web-based interface. Firstly, we obtained informed consent, and participants completed a brief, optional demographic questionnaire. We explained how the task would be solved and then presented the questions accordingly. They had to complete each task to proceed to the next one. In each task, participants compared the two clips on six traits: *openness*, *conscientiousness*, *extraversion*, *agreeableness*, *emotional stability*, and *naturalness*. Each question has a forced-choice format with three possible choices: *Left*, *Equal*, or *Right*. We included naturalness as a control dimension, allowing us to perform a validity check. This check allows us to verify that observed personality differences are driven by the intended choice of participant, rather than being confounded by technical rendering artifacts or motion instability inherent in neural talking-head synthesis pipelines.

Data Processing and Quality Control. Responses were encoded numerically as -1 (left video perceived stronger), 0 (both videos perceived equal), and $+1$ (right video perceived stronger). These scores were aggregated per comparison pair and trait across N=80 participants for further analysis. The results are then divided into the same-expression, different-head-orientation group, and the different-expression pair comparisons. The experiment interface required participants to complete all tasks before participating in the study, and their progress was automatically saved to allow recovery in the event of an interruption. Upon completing the final pair, participants were automatically redirected to Prolific using a completion code, ensuring reliable submission tracking and compensation.

5 Analysis and results

Pairwise Expression Comparison. To estimate the relative preference strength of each facial expression for each personality trait, we fitted a Bradley–Terry model with the Davidson [37] extension, which explicitly incorporates tied responses (“equal”) via a tie parameter δ. This model yields a worth parameter π_i for each emotion, where a higher worth indicates a stronger association with the target trait. Worth parameters are normalized to sum to 1 within each trait (chance level = 0.25). The estimated worths are visualized in Fig. 2 and reported in Table 1.

Because each participant evaluated all six emotion pairs, responses are not independent. To obtain valid inferential statistics under this repeated-measures design, we fitted a BT model as a binomial Generalized Estimating Equation (GEE) with an exchangeable working correlation structure. This model adjusts standard errors for within-participant dependence. In the GEE formulation, tied responses are excluded from the binary likelihood and captured descriptively by the Davidson tie parameter. All p-values were corrected using the Benjamini–Hochberg (BH) procedure. Omnibus Wald tests evaluate whether emotions differ in perceived trait strength, and coefficient-level effects were parameterized using sum-to-zero (effect coding), so no single emotion served as a reference category; coefficients indicate deviations from the grand mean across emotions. All six omnibus tests were statistically significant (Table 1; all $p_{\mathrm{BH}}<.001$), indicating that facial expression influences perceived personality across all evaluated traits.

The Bradley–Terry–Davidson worth estimates (Fig. 2) showed that happy expressions received the highest worth for five of the six traits: extraversion (π=.71), agreeableness (π=.99), emotional stability (π=.56), openness (π=.78), and naturalness (π=.69). These patterns were consistent with the GEE-based omnibus tests, confirming that perceived personality ratings differed significantly across expressions.

Angry expressions showed a distinctive pattern. As shown in Fig. 2, they were strongly associated with extraversion (π=.28), while being consistently rated lowest on agreeableness ($\pi\approx$0), emotional stability (π=.01), and openness (π=.01). This pattern suggests that high-arousal expressions may be perceived as socially dominant but less prosocial or emotionally stable. These findings were consistent with the GEE-based omnibus tests.

Conscientiousness showed the weakest differentiation across expressions (Fig. 2; Table 1). Surprised (π=.39), sad (π=.30), and happy (π=.25) received similar worth values, indicating limited separation between expressions. These patterns were consistent with the GEE-based omnibus tests, suggesting that conscientiousness is less readily inferred from brief facial-expression cues.

Sad and surprised expressions occupied similar middle positions across several traits (Fig. 2). Their value weights were especially similar for openness, emotional stability, and naturalness, indicating minimal differentiation between the two expressions. These patterns were consistent with the GEE-based omnibus tests. This convergence suggests that these two expressions convey overlapping personality signals.

The Davidson tie parameter δ (Table 1) provides additional insight into trait discriminability. The observed tie rate varied substantially across traits: naturalness had the highest rate (38.8%), followed by openness (25.6%) and conscientious-

Table 1 Bradley–Terry–Davidson worth parameters per emotion and trait (E: Extraversion, A: Agreeableness, C: Conscientiousness, ES: Emotional Stability, O: Openness, N: Naturalness). Worth values are on the probability scale (sum to 1 per trait; chance=0.25). The Davidson tie parameter δ captures the propensity for tied ("equal") responses; $\delta > 1$ indicates more ties than standard BT predicts. Omnibus Wald χ^2 tests (BH-corrected) from the BT–GEE (Generalized Estimating Equation (GEE) model confirm significant emotion effects for all traits. Tie % gives the descriptive rate of "equal" responses

Trait	Happy	Angry	Sad	Surprised	δ	Tie %	$\chi^2(3)$	p_{BH}
E	**0.714**	0.280	0.003	0.003	0.83	13.8%	74.0	$<.001^{***}$
A	**0.993**	0.000	0.003	0.004	1.50	16.7%	89.5	$<.001^{***}$
C	0.254	0.058	0.302	**0.386**	0.71	23.3%	32.6	$<.001^{***}$
ES	**0.562**	0.006	0.201	0.231	0.55	14.0%	89.2	$<.001^{***}$
O	**0.777**	0.013	0.103	0.108	1.13	25.6%	79.4	$<.001^{***}$
N	**0.689**	0.093	0.092	0.125	1.54	38.8%	51.9	$<.001^{***}$

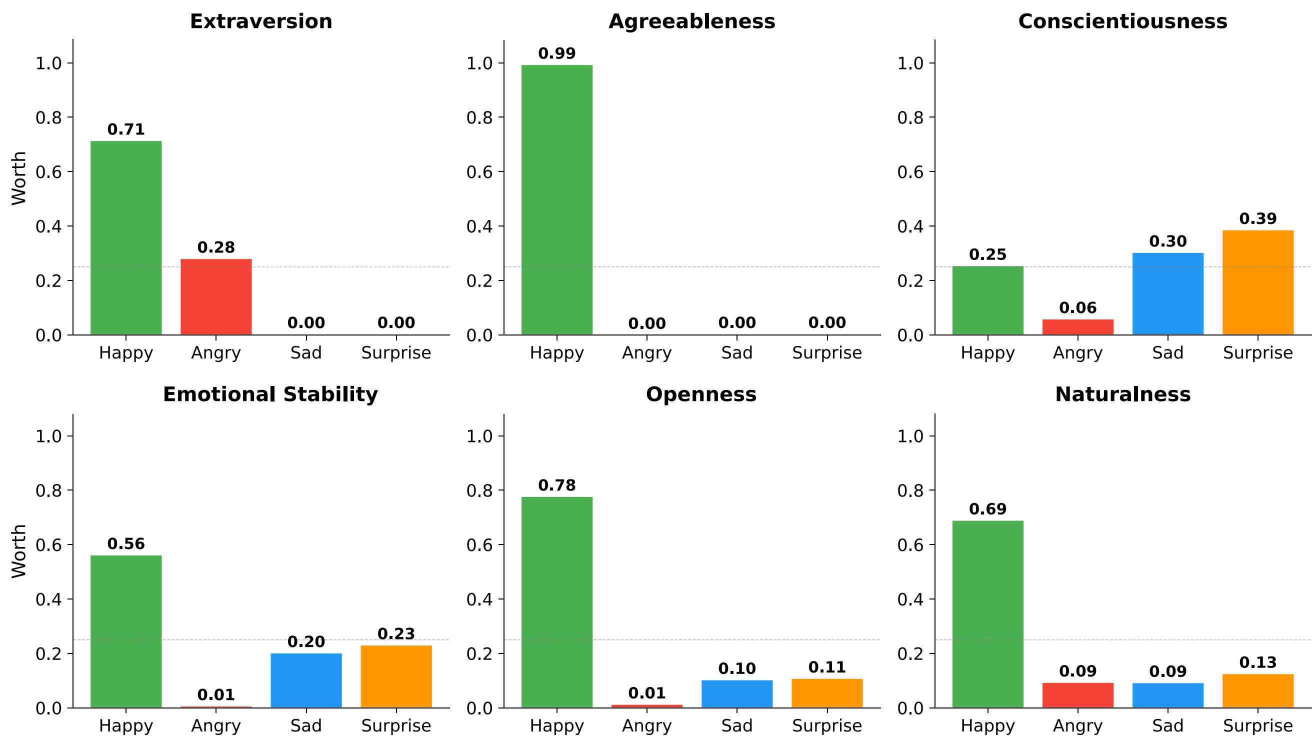

Fig. 2 Bradley–Terry–Davidson worth parameters for perceived personality traits across facial expressions. Each bar represents the estimated worth of an emotion for a given trait (sum to 1 per trait). The dashed line indicates the chance level (0.25). Ties are modeled via the Davidson δ parameter (see Table 1)

ness (23.3%), while extraversion (13.8%), emotional stability (14.0%), and agreeableness (16.7%) showed lower tie rates.

These differences reflect both the model's tie propensity parameter δ and the spacing of worth values: traits with closely spaced worths (e.g., conscientiousness) produce more ties, whereas traits with strongly separated worths (e.g., agreeableness) produce fewer ties even when δ is high. The overall tie rate was 22.0%, indicating moderate ambiguity in participants' judgments.

Influence of Head Orientation. To evaluate whether head orientation (left-facing profile vs. frontal) affects personality perception, we fitted a cumulative link mixed model (CLMM) for each trait, with emotion as a fixed effect and participant as a random intercept to account for the repeated-measures design. The response categories were modeled as an ordinal variable with the ordering *left-facing* < *equal* < *frontal*. Emotion effects were parameterized using sum-to-zero (effect coding), such that no single emotion serves as a reference category. Coefficients represent deviations from the grand mean across emotions. Results are given in Table 2 and Fig. 3.

Across all traits and emotions, the equal response dominated. Predicted probabilities from the CLMM (Table 2c), together with the coefficient patterns shown in Fig. 3, indi-

Table 2 Cumulative Link Mixed Model (CLMM) results for head orientation preference. The omnibus likelihood-ratio test (LRT) assesses whether head orientation preference differs across emotions for each trait. Fixed effects represent deviations from the grand mean across all emotions (sum-to-zero coding); positive estimates indicate a shift toward frontal preference. Predicted probabilities are reported at the average participant level (random effect=0). All p-values are BH-corrected (n.s.: not significant)

(a) Omnibus LRT: emotion effect on orientation preference

Trait	$\chi^2(3)$	p_{BH}	Interpretation
Extraversion	10.4	.047*	Happy → more frontal
Agreeableness	2.5	.481	n.s
Conscientiousness	2.6	.481	n.s
Emotional Stability	4.4	.334	n.s
Openness	4.9	.334	n.s
Naturalness	18.3	.002**	Surprise → less frontal
Naturalness	18.3	.002**	Anger ← more frontal

(b) Significant fixed effects (deviation from grand mean)

Trait	Emotion	$\hat{\beta}$ (SE)	p_{BH}
Extraversion	Happy	0.620 (0.201)	.024*
Naturalness	Angry	0.607 (0.206)	.026*
Naturalness	Surprise	−0.640 (0.201)	.024*

(c) Predicted response probabilities (average participant)

Trait	Emotion	P(left)	P(equal)	P(frontal)
Extraversion	Angry	.117	.689	.194
	Happy	.046	.558	.396
	Sad	.092	.669	.239
	Surprise	.093	.670	.237
Naturalness	Angry	.071	.580	.349
	Happy	.090	.618	.292
	Sad	.159	.664	.177
	Surprise	.209	.658	.134

Fixed effects are reported using sum-to-zero (effect coding); coefficients indicate deviations from the grand mean across emotions. Positive values indicate a greater frontal preference; negative values indicate a greater left-facing preference. $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$ (BH-corrected)

cate that P(equal) consistently exceeds both left and frontal choices, ranging from approximately.56 to.75 across conditions. This convergence of evidence shows that head orientation has a limited overall effect on personality perception, with participants most often reporting no perceptual difference between orientations.

Only extraversion and naturalness showed significant omnibus effects (Table 2a). For extraversion ($\chi^2(3)$=10.4, p_{BH}=.047), happy expressions showed a significantly stronger frontal preference (positive deviation from the grand mean; $\hat{\beta}$=0.62, p_{BH}=.024). As illustrated in Fig. 3, this corresponds to a clear directional shift toward frontal preference for happy, with the predicted frontal probability increasing from.19 to.40 across conditions (see Table 2), despite the equal category remaining the most frequent response. The remaining traits—agreeableness, conscientiousness, emotional stability, and openness—did not show significant effects (all $p_{\mathrm{BH}}\geq$.334), indicating that head orientation does not meaningfully influence perceptions of these traits.

6 Discussion

Our findings indicate that facial expressions significantly influence personality perception, supporting H1 for RQ1. Happy expressions yielded the strongest signal for socially positive traits, consistently receiving the highest attribution scores to extraversion, agreeableness, openness, and emotional stability. Angry expressions were linked to lower agreeableness and emotional stability, as expected, but also served as a secondary signal for high extraversion, suggesting that high-arousal negative emotions can still convey social dominance. Conversely, conscientiousness appeared less susceptible to isolated facial cues in our short-form neural videos, indicating it may be a more latent trait that requires longer exposure or multi-modal context, such as speech or specific behavioral patterns, to be reliably inferred. These results support H1, confirming that facial expressions are a primary, trait-dependent driver of personality perception. All associations are consistent with previous findings on personality and facial expression parameters, in which extraversion, agreeableness, and expressions of happiness are positively

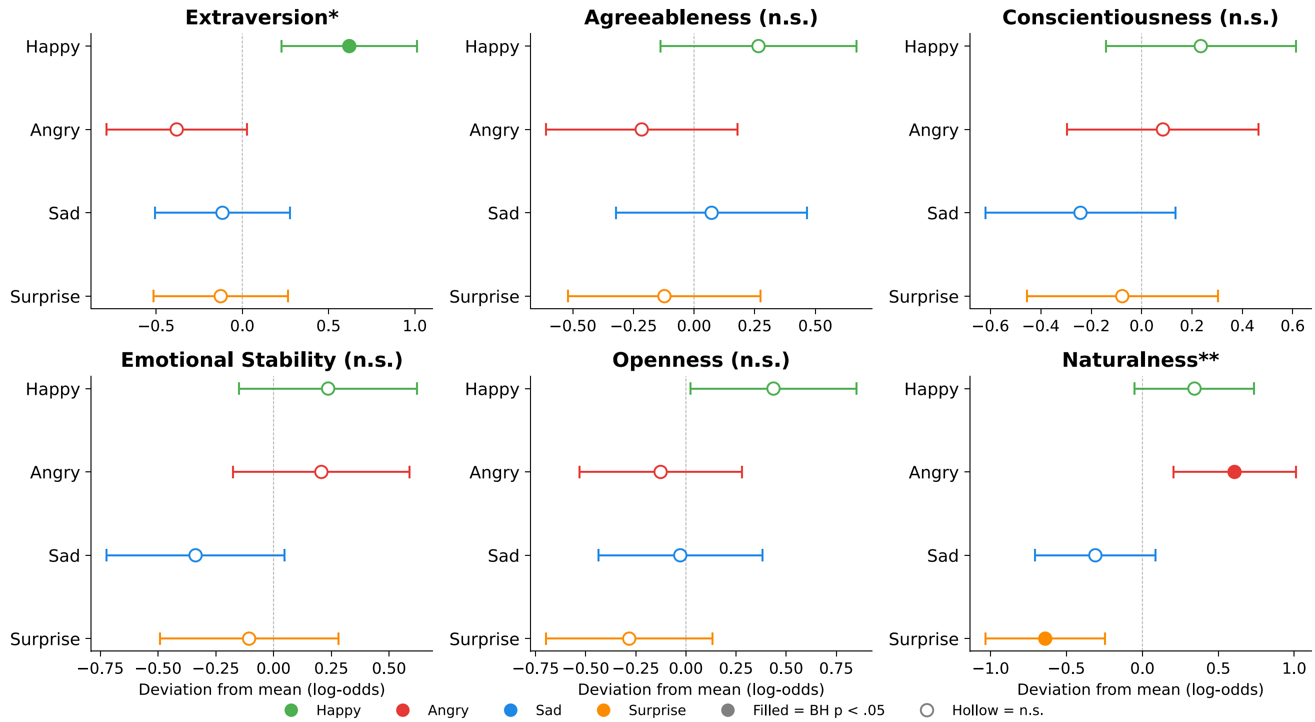

Fig. 3 CLMM emotion coefficients ($\hat{\beta}$) for head orientation preference per trait, with 95% confidence intervals. Coefficients are estimated using sum-to-zero (effect coding) and represent deviations from the grand mean across all emotions. Positive values indicate a shift toward frontal preference, negative values indicate a shift toward left-facing preference. Filled markers denote significant deviations from the mean (BH-corrected $p<.05$). Asterisks in panel titles indicate significant omnibus LRT effects ($^{*}p < .05$, $^{**}p < .01$). n.s.: not significant

correlated, while conscientiousness is associated with the lowest level of expressivity [28, 39].

Regarding RQ2, our results reveal that head orientation lacks the perceptual salience of facial expressions and has a limited overall effect on personality impressions. The dominance of the “equal” response category suggests that orientation is often perceived as a neutral parameter in short-form neural video. Nevertheless, our findings show that orientation can serve as a subtle cue for extraversion, with frontal views significantly increasing the perceived association with happy expressions. However, these effects did not generalize to agreeableness, conscientiousness, emotional stability, or openness; thus, H2 is only weakly supported. Although head orientation alone is not a dominant cue, it can subtly shape perception when interacting with specific facial expressions. For instance, it can be strategically adjusted to control perceived social engagement.

There were slight associations between naturalness and anger in frontal views, and naturalness and surprise in left-facing views. These results indicate that angry expressions were perceived as more natural when the character directly faces the camera, and surprised expressions were perceived as more natural when the character looks to the side. The association between frontal orientation and increased naturalness for angry expressions supports the Shared Signal Hypothesis [40]. From an evolutionary perspective, anger is a confrontational signal. Thus, a frontal orientation provides a clear signal of intent, whereas an averted view may dilute the perceived authenticity of the threat. In contrast, surprise is a reactive and information-gathering emotion. Unlike anger, which is directed at someone, surprise is usually triggered by an environmental stimulus. In our videos, when the character faces the side, it implies looking at an object or event outside the frame, contextualizing the association with surprise.

For RQ3, we observe that effect sizes vary across personality traits, with extraversion and agreeableness exhibiting the largest effect sizes, followed by openness. In contrast, we found conscientiousness to be relatively insensitive, consistent with the literature [28, 39]. Additionally, sad and surprised expressions were perceived as highly similar across multiple traits, indicating overlapping perceptual signals in brief visual stimuli. Overall, the analysis highlights that not all traits are equally inferable from facial cues, partially supporting H3.

Beyond technical realism, integrating neural-rendered virtual humans into daily life requires addressing ethical concerns such as perceptual bias and behavioral transparency [35]. By quantifying the influence of subtle nonverbal cues on personality attribution, our study provides an empirical basis for ensuring fairness and preventing unintended manipulation in the design of digital humans.

These results offer practical insights for designing virtual agents and digital humans. Visual cues shape perceived personality traits, but their effects depend on how they are combined. Frontal views improve naturalness most for angry expressions, while surprised and sad expressions show weaker frontal preferences. To convey extraversion, designers should pair high-energy positive expressions with a frontal view. For agreeableness, expression matters more than head orientation. Traits such as conscientiousness may require support from auditory cues or other modalities.

A limitation of this study is the use of a single identity. Although fixing identity allowed us to isolate the perceptual effects of facial expression and head orientation under controlled conditions, the observed results may still be partly identity-specific. Static properties of the selected face, such as facial morphology, gender presentation, and ethnic appearance, may have influenced participants' personality judgments independently of the manipulated motion cues. Therefore, the present findings should be interpreted as evidence obtained under a controlled single-identity setting rather than as fully generalizable across identities.

7 Conclusion and future work

This work focuses on analyzing perceived personality in neural-rendered talking head videos. We designed a user study with eight video clips featuring combinations of four facial expressions and two head orientations. The analysis of study results suggests that facial expressions influence the perception of personality, with happy expressions leading to higher perceived levels of extraversion, agreeableness, and emotional stability than other emotions. Sad and surprised expressions yielded highly similar trait attribution profiles, suggesting that these distinct emotional states may convey overlapping social signals in brief, neural-rendered stimuli. Findings further indicate that head orientation has a limited overall effect on personality perception, as reflected by the high proportion of "equal" responses across conditions. However, subtle directional effects emerged in specific cases. For example, frontal views increased perceived extraversion for happy expressions, suggesting that head orientation can modulate perception when combined with facial expression. The naturalness judgments show that participants were not evaluating the stimuli solely in terms of rendering realism, but were sensitive to the intended changes in facial expression and head orientation.

Our findings have practical implications for the design of virtual agents, digital humans, and neural-rendered avatars in several application domains. In entertainment applications such as computer animation, games, and virtual storytelling, the results can help animators control how characters are perceived through subtle combinations of facial expression and head orientation. In human-computer interaction, the findings can support the design of embodied conversational agents, virtual assistants, and telepresence avatars whose nonverbal behavior is adapted to convey intended social traits more consistently. The results may also be useful in social robotics, virtual training, and educational simulations, where first impressions and perceived personality can influence user trust, comfort, and engagement.

Future work can extend this framework in several directions. First, auditory channels should be studied together with visual cues, particularly for traits such as conscientiousness that appear less sensitive to short visual signals alone. Second, the current study focuses on short-term impressions, and longer interaction settings may reveal whether these effects remain stable over time. Finally, future research should test multiple identities and broader demographic variation to evaluate the generalizability of the observed effects.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11760-026-05442-y.

Author Contributions M.K. designed and implemented the framework and performed the experimental study. U.G. and S.S. conceptualized and defined the methodology. A.Ü.E. contributed to the implementation and experimental study. F.D. contributed to the interpretation and analysis of data from the user study experiments. U.G. supervised the study. All authors contributed to the writing and reviewing.

Funding The authors received no financial support for the research, authorship, and/or publication of this article.

Data Availability The neural-rendered talking head videos generated from the Multi-PIE dataset are publicly available at https://github.com/Melik3Kara/Neural-Rendered-Talking-Head

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Ethics approval Bilkent University Ethical Committee for Human Research approved the study with Decision Number 781 at the meeting İAEK_2025_12_07_01.

References

1. Willis, J., Todorov, A.: First impressions: Making up your mind after a 100-ms exposure to a face. Psychol. Sci. **17**(7), 592–598 (2006). https://doi.org/10.1111/j.1467-9280.2006.01750.x
2. Ambady, N., Rosenthal, R.: Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. Psychol. Bull. **111**, 256–274 (1992)
3. Todorov, A., Said, C.P., Engell, A.D., Oosterhof, N.N.: Understanding evaluation of faces on social dimensions. Trends Cogn.

Sci. **12**(12), 455–460 (2008). https://doi.org/10.1016/j.tics.2008.10.001
4. Mignault, A., Chaudhuri, A.: The many faces of a neutral face: Head tilt and perception of dominance and emotion. J. Nonverbal Behav. **27**(2), 111–132 (2003). https://doi.org/10.1023/A:1023914509763
5. McCloud, S.: Understanding Comics: The Invisible Art. Tundra Publishers, Toronto, Canada (1993)
6. Manning, A.D.: Understanding comics: The invisible art. IEEE Trans. Prof. Commun. **41**(1), 66–69 (1998). https://doi.org/10.1109/TPC.1998.661632
7. Thomas, F., Johnston, O.: The Illusion of Life: Disney Animation. Disney Editions Deluxe. Disney Books, Westminster, MD, USA (1995)
8. Siarohin, A., Lathuilière, S., Tulyakov, S., Ricci, E., Sebe, N.: First order motion model for image animation. In: Proceedings of the 33rd International Conference on Neural Information Processing Systems. NeurIPS '19, (pp. 7137–7147). Curran Associates Inc, Red Hook, NY, USA (2019)
9. Prajwal, K.R., Mukhopadhyay, R., Namboodiri, V.P., Jawahar, C.V.: A lip sync expert is all you need for speech to lip generation in the wild. In: Proceedings of the 28th ACM International Conference on Multimedia. MM '20, (pp. 484–492). ACM, New York, NY, USA (2020). https://doi.org/10.1145/3394171.3413532
10. Guo, J., Zhang, D., Liu, X., Zhong, Z., Zhang, Y., Wan, P., Zhang, D.: LivePortrait: Efficient Portrait Animation with Stitching and Retargeting Control. arxiv.org/abs/2407.03168 (2025)
11. Tewari, A., Thies, J., Mildenhall, B., Srinivasan, P., Tretschk, E., Yifan, W., Lassner, C., Sitzmann, V., Martin-Brualla, R., Lombardi, S., Simon, T., Theobalt, C., Nießner, M., Barron, J.T., Wetzstein, G., Zollhöfer, M., Golyanik, V.: Advances in neural rendering. Computer Graphics Forum **41**(2), 703–735 (2022). https://doi.org/10.1111/cgf.14507
12. Chen, J., Wang, F., Huang, Z., Zhou, Q., Li, K., Guo, D., Zhang, L., Yang, X.: Towards seamless interaction: Causal turn-Level modeling of interactive 3D conversational head dynamics (2025). arxiv.org/abs/2512.15340
13. Thies, J., Zollhöfer, M., Theobalt, C., Stamminger, M., Niessner, M.: Headon: real-time reenactment of human portrait videos. ACM Transactions on Graphics **37**(4), 164 (2018). https://doi.org/10.1145/3197517.3201350
14. Little, A.C., Perrett, D.I.: Using composite images to assess accuracy in personality attribution to faces. Br. J. Psychol. **98**(1), 111–126 (2007). https://doi.org/10.1348/000712606X109648
15. Biel, J.-I., Teijeiro-Mosquera, L., Gatica-Perez, D.: FaceTube: predicting personality from facial expressions of emotion in online conversational video. In: Proceedings of the 14th ACM International Conference on Multimodal Interaction. ICMI '12, (pp. 53–56). ACM, New York, NY, USA (2012). https://doi.org/10.1145/2388676.2388689
16. Sonlu, S., Doğan, Y., Ergüzen, A.Ü., Ünalan, M.E., Demirci, S., Durupinar, F., Güdükbay, U.: Towards understanding personality expression via body motion. In: Workshop on Multi-modal Affective and Social Behavior Analysis and Synthesis in Extended Reality. MASSXR '24, pp. 628–631. IEEE, Piscataway, NJ, USA (2024)
17. Todorov, A., Porter, J.M.: Misleading first impressions: Different for different facial images of the same person. Psychol. Sci. **25**(7), 1404–1417 (2014). https://doi.org/10.1177/0956797614532474
18. Livingstone, S.R., Russo, F.A.: The Ryerson Audio-Visual Database of Emotional Speech and Song (RAVDESS): A dynamic, multimodal set of facial and vocal expressions in North American English. PLoS ONE **13**(5), 0196391 (2018). https://doi.org/10.1371/journal.pone.0196391
19. Thomas, S., Ferstl, Y., McDonnell, R., Ennis, C.: Investigating how speech and animation realism influence the perceived personality of virtual characters and agents. In: Proceedings of the IEEE Conference on Virtual Reality and 3D User Interfaces. VR '22, (pp. 11–20). IEEE, Piscataway, NJ, USA (2022)
20. Gross, R., Matthews, I., Cohn, J., Kanade, T., Baker, S.: Multi-PIE Face Dataset. https://www.kaggle.com/datasets/aliates/multi-pie (2010)
21. Gross, R., Matthews, I., Cohn, J., Kanade, T., Baker, S.: Multi-PIE. Image Vis. Comput. **28**(5), 807–813 (2010). https://doi.org/10.1016/j.imavis.2009.08.002
22. Bruce, V., Young, A.: Face Perception. Psychology Press, London, U.K. (2012)
23. Wilson, H.R., Wilkinson, F., Lin, L.-M., Castillo, M.: Perception of head orientation. Vision. Res. **40**(5), 459–472 (2000). https://doi.org/10.1016/S0042-6989(99)00195-9
24. Sonlu, S., Bendiksen, B., Durupinar, F., Güdükbay, U.: Effects of embodiment and personality in LLM-based conversational agents. In: Proceedings of the IEEE Conference Virtual Reality and 3D User Interfaces. VR '25, (pp. 718–728). IEEE, Piscataway, NJ, USA (2025)
25. Aylett, M.P., Vinciarelli, A., Wester, M.: Speech synthesis for the generation of artificial personality. IEEE Trans. Affect. Comput. **11**(2), 361–372 (2020)
26. Ijuin, K., Jokinen, K.: Exploring gaze behaviour and perceived personality traits. In: Meiselwitz, G. (ed.) Social Computing and Social Media. Design, Ethics, User Behavior, and Social Network Analysis, (pp. 504–512). Springer, Cham (2020)
27. Ruhland, K., Peters, C.E., Andrist, S., Badler, J.B., Badler, N.I., Gleicher, M., Mutlu, B., McDonnell, R.: A review of eye gaze in virtual agents, social robotics and HCI: Behaviour generation, user interaction and perception. Computer Graphics Forum **34**(6), 299–326 (2015). https://doi.org/10.1111/cgf.12603
28. Kodavati, S., Mei, W.A.H., Kanade, A., Durupinar, F.: A personality-labeled semantic dataset from facial expressions, gaze, and head movement cues. Computers & Graphics **133**, 104455 (2025). https://doi.org/10.1016/j.cag.2025.104455
29. Daoudi, M., Srivastava, A., Veltkamp, R.: 3D Face Modeling. Analysis and Recognition. John Wiley & Sons, Hoboken, NJ, USA (2013)
30. Weier, M., Stengel, M., Roth, T., Didyk, P., Eisemann, E., Eisemann, M., Grogorick, S., Hinkenjann, A., Kruijff, E., Magnor, M., Myszkowski, K., Slusallek, P.: Perception-driven accelerated rendering. Computer Graphics Forum **36**(2), 611–643 (2017). https://doi.org/10.1111/cgf.13150
31. Rekik, R., Wuhrer, S., Hoyet, L., Zibrek, K., Olivier, A.-H.: A survey on realistic virtual human animations: Theory and practice. Computer Graphics Forum **43**(2), 15033 (2024). https://doi.org/10.1111/cgf.15064
32. Geller, T.: Overcoming the uncanny valley. IEEE Comput. Graphics Appl. **28**(4), 11–17 (2008)
33. Gross, R., Matthews, I., Cohn, J., Kanade, T., Baker, S.: Multi-PIE. In: Proceedings of the 8th IEEE International Conference on Automatic Face & Gesture Recognition. FG '08, (pp. 1–8) (2008). https://doi.org/10.1109/AFGR.2008.4813399
34. Ekman, P., Friesen, W.: Constants across cultures in the face and emotion. J. Pers. Soc. Psychol. **17**, 124–129 (1971). https://doi.org/10.1037/h0030377
35. Knob, P., Pinho, G., Silva, G.F., Montanha, R., Peres, V., Araujo, V., Musse, S.R.: Surveying the evolution of virtual humans expressiveness toward real humans. Computers & Graphics **123**, 104034 (2024). https://doi.org/10.1016/j.cag.2024.104034
36. Zibrek, K., Kokkinara, E., Mcdonnell, R.: The effect of realistic appearance of virtual characters in immersive environments - does the character's personality play a role? IEEE Trans. Visual Comput. Graphics **24**(4), 1681–1690 (2018). https://doi.org/10.1109/TVCG.2018.2794638

37. Davidson, R.R.: On extending the bradley-terry model to accommodate ties in paired comparison experiments. J. Am. Stat. Assoc. **65**(329), 317–328 (1970)
38. Whitehead, J.: Sample size calculations for ordered categorical data. Stat. Med. **12**(24), 2257–2271 (1993). https://doi.org/10.1002/sim.4780122404
39. Sonlu, S., Güdükbay, U., Durupinar, F.: A conversational agent framework with multi-modal personality expression. ACM Transactions on Graphics **40**(1), 7 (2021)
40. Adams, R.B., Jr., Kleck, R.E.: Perceived gaze direction and the processing of facial displays of emotion. Psychol. Sci. **14**(6), 644–647 (2003)